University of Bath

PHD

A Statistical Analysis of Engagement in Arabic Language MOOCs

Almansour, Shahad

Award date:
2020

Awarding institution:
University of Bath

Link to publication

Alternative formats
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

Take down policy
If you consider content within Bath's Research Portal to be in breach of UK law, please contact: openaccess@bath.ac.uk with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

Download date: 05. May. 2023

A Statistical Analysis of Engagement in Arabic Language MOOCs

submitted by

Shahad M Almansour

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Computer Sciences

September 2019

COPYRIGHT

Attention is drawn to the fact that copyright of this thesis rests with its author. This copy of the thesis has been supplied on the condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the prior written consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation.

Signature of Author ....

Shahad M Almansour

To Khaled

I would never have become who I am today if you weren't by my side motivating and inspiring me.

Acknowledgement

I would like to take this opportunity to offer my sincerest gratitude to my lead supervisor, Dr. Alan Hayes, for all his guidance, encouragement and advice. He was very understanding and caring about all the circumstances that I faced as an international student.

My sincerest gratitude to my second supervisor, Dr. John Power, who was equally understanding and caring. He provided me with support and guidance throughout my PhD and always found the time in his busy schedule to offer me valuable feedback. I was extremely lucky to work under these supportive supervisors.

Thank you to Edraak research team for their cooperation in this research. Many thanks to everyone at the Mathematics Resources Centre (MASH) at the University of Bath for helping me and many other students with maths and statistics problems. Thank you to everyone in the Computer Science Department at the University of Bath for the comfortable and friendly environment, and for always being supportive of each other.

My deepest gratitude to my family: my parents Mousa Almansour and Latifah Alomayrah, my brother Ali, my sisters Maram, Tasnim, Banan and Raghad for their encouragement and unconditional love and support. Special thanks to Latifah Alshammary, Meriem Bouazzaoui, Nour Alhusein, Hanan Alkandari and Fahid Mohammed for being such supportive friends.

Finally, I would like to thank my sponsor, the Ministry of Education in Saudi Arabia, for their funding and support. Similarly, thank you to the Saudi Arabia Cultural Bureau in London for taking care of me and many other students during our study abroad.

September, 2019
Bath, UK

Abstract

As massive open online courses (MOOCs) gain popularity as an educational resource with minimal time, space and fee restrictions, many researchers invested in studying the engagement of learners with MOOCs. Such studies have enabled the identification of learners' general characteristics and engagement patterns, which can help the MOOC's providers to better understand the learners' needs and accordingly enhance their learning experiences. In 2013, MOOCs reached the Arabic region when two MOOC's platforms launched, Rwaq (in Saudi Arabia) and Edraak (in Jordon). They are now considered the leading Arabic MOOC platforms with more than 400 courses and over 70,000 daily visitors, combined. However, Arabic MOOC providers are only focusing on launching new platforms and delivering MOOCs that cover many subjects. The field of MOOC research lacks investigation of the adoption of MOOCs in the Arabic region. This might prevent the improvement of the learning experience of the Arabic-speaking learners. This research aims to study the adoption of MOOCs in the Arabic-speaking region. This can be achieved by: (1) identifying Arabic-speaking learners' characteristics, (2) analysing their engagement with MOOCs' contents, and (3) identifying differences and similarities in learner's engagement between Arabic- and English-speaking learners.

As the subjects of many studies, the leading MOOC platforms in the Western world have a good understanding of their learners and their needs. Following their lead, this research uses two fundamental works in the area of learner engagement with MOOCs. These are: Kizilcec et al. [1] and Ferguson and Clow [2] who analysed learner engagement with MOOCs in the Coursera and Futurelearn platforms, respectively. Kizilcec et al. [1] introduced a classification method to classify learners based on their weekly interaction with the MOOC contents. They used two variables to compute learners' weekly score, videos and assessments. Then they used these scores in a one-dimensional K-means clustering algorithm, which clustered their learners into four groups. This classification method was modified by Ferguson and Clow [2] and applied to the Futurelearn platform. They included one more variable to compute learners' weekly score, weekly comments. Then they used these scores in a multidimensional K-means clustering algorithm to prevent losing useful data when composing the weekly scores into a singular digit for the one-dimensional K-means clustering algorithm. Their classifica-

tion method clustered their learners into seven groups, two of which are found in Kizilcec et al.'s result. We applied both classification methods (one- and multidimensional K-means clustering algorithm) to our data that we obtained from the Edraak platform. We found that the learner engagement results were similar. Each classification method produced the same three groups that represent Edraak's learners' engagement types, which are the following:

- Sampling: represents learners who had no interaction with the MOOC contents or dropped out of the MOOC after the first week.
- Disengaging: represents learners who had a good interaction with the MOOC contents at the beginning of the MOOC, but had dropped out by the middle of the course.
- Completing: represents learners who had a good interaction with the MOOC contents, and completed the MOOC.

Comparing our results with Kizilcec et al.'s [1] and Ferguson and Clow's [2] showed that the engagement of Edraak's learners with MOOCs is closer to Coursera's learners than Futurelearn's learners. Edraak's and Coursera's learners have a lower completion rate than Futurelearn's learners, in addition to the poor use of discussion forum in Edraak's and Coursera's platforms compared to Futurelearn's platform. Moreover, considering pedagogical approaches adopted by these platforms, we found that the absence of a clear pedagogical approach had a negative effect on the engagement of Edraak's learners.

Contents

List of Figures

List of Tables

Introduction

Massive Open Online Courses (MOOC) have proliferated as a distance education option. They have become increasingly popular with the increase in MOOC providers and learners since 2008, when the first MOOC was launched [4, 5, 6]. Given MOOCs' open access, flexible learning time and space, free-of-charge materials and large-scale interaction with learners and educators around the world [7], some researchers have found MOOCs to be supportive of traditional education [8, 9]. Ferguson et al. [10] believed that MOOCs promote the understanding of how to deliver education online to a large number of learners. They predicted that by the year 2030, top universities would have a large, worldwide community of learners, and that universities role would be researching ways to improve learners' experience [10].

With the popularity of MOOCs in the Western world, some researchers have investigated the use of MOOCs in developing countries. Warusavitarana et al. [11] studied the use of MOOCs in Sri Lankan institutions of higher education. They asked several researchers to participate in various MOOC platforms and report their experience. They found MOOCs to have numerous benefits for learning without the restrictions of time and place and allow access to the most recent educational resources [11]. Abas [12] examined the use of MOOCs in southeast Asia, especially Malaysia and Indonesia, and reported that the growing demand for access to higher education in these countries has led to increased interest in MOOCs. Venkataraman and Kanwar [13], from the Commonwealth of Learning (COL) [1], designed a MOOC in partnership with the Indian Institute of Technology Kanpur to support human development. Their MOOC addressed ways in which mobile devices can aid in areas of development (e.g., agriculture). They successfully attracted over 2,000 participants, 90% of whom were from

[1] The COL is an intergovernmental organization established in Vancouver 1987, aiming to improve access to quality education and training in developing countries.

developing countries [13]. Venkataraman and Kanwar [13] expected this use of MOOCs to grow as government agencies showed interest in using MOOCs to improve literacy, health and economic standing.

The increased popularity of MOOCs around the world reached the Arab region in 2013 with the launch of two Arabic-language platforms, Edraak and Rwaq [14, 15]. Although these platforms are designed and deliver their MOOCs in Arabic, their structure and methods of delivering materials are similar to some of the leading MOOC platforms (e.g., edX and Coursera) in that they provide their learners with video lectures, weekly assessments, articles, quizzes and a discussion forum. The leading MOOC platforms provided researchers with data to study their learners in an effort to gain a better understanding of the learners' needs and design better MOOCs accordingly. In 2014, Pang et al. [16] studied MOOC data that were provided from three leading MOOC platforms (Coursera, Udacity and edX). They linked the high enrolment in Coursera, as compared to edX and Udacity, with the provision of MOOCs in multiple languages; this showed the importance of localising MOOCs by using local languages. Kizilcec et al. [1] investigated learners' engagement patterns in Coursera, by classifying the learners into groups based on learner interaction with the MOOCs' contents (i.e. videos and assessments). Ferguson and Clow [2] followed Kizilcec et al.'s approach and performed a similar study to classify learners on the Futurelearn platform, where they identified more patterns of engagement. These studies demonstrate the importance of analysing MOOC learner data to understand learner needs; this contributes to the improvement of MOOC design. Unfortunately, Arabic MOOC platforms lack such studies to understand the needs of Arabic learners and their objectives in using MOOCs.

This study investigated the adoption of MOOCs in Arabic-speaking countries. We compared and contrasted the engagement of MOOC learners in Arabic and in English to identify similarities and differences between the two mediums. We analysed learners' data in terms of completion rate, dropout rate, gender proportion, number of comments and educational qualification. This examination identifies the characteristics of Arabic-language MOOC learners and provides a guide for MOOC designers and providers to improve their platforms based on learners' needs.

Objectives and research questions

This study aimed to improve the understanding of Arabic learners' use of MOOCs. This understanding can be obtained by identifying the learners' characteristics, which helped in specifying the targeted sample. To do so, we obtained MOOC data from the Arabic MOOC platform Edraak. A MOOC entitled Java Programming 1, presented by the Arab Open University in Jordan, was provided to us by the Edraak team. This MOOC ran for six weeks and targeted Arabic learners at all levels. The research aimed to cluster learners from this MOOC based on their interaction with the MOOCs' contents (videos and assessments). This clustering was adopted from Kizilcec et al. [1] and Ferguson and Clow [2], who studied learner engagement with MOOCs on two leading platforms, Coursera and Futurelearn, respectively.

Moreover, this research determined the most applicable K-means clustering approach to our data to understand Edraak learners' engagement with MOOCs. In addition, to further clarify our understanding of Arabic-speaking MOOC learners, this research aimed to compare our data from the Edraak platform with the leading MOOC platforms, Coursera and Futurelearn. Identifying similarities and differences among learners from different platforms can help in sharing and adopting ideas that can enhance the MOOC experience. This research provides information and tools that can be used by Arabic-speaking platform providers and designers that can guide future Arabic MOOC research. To achieve these aims, our research attempted to answer the following questions:

Q1. Arabic-speaking learners profile: Who are the users of the Arabic MOOC platforms?

Q2. Engagement types of Arabic-speaking learners: What are the engagement types of the Arabic-speaking learners?

Q3. MOOC engagement between Arabic- and English-speaking learners How is the engagement of Arabic-speaking learners in a MOOC similar to or different from the engagement of English-speaking learners in a MOOC?

Organisation of the thesis

Chapter 1 introduces the various types of MOOCs and describes their presence in the Arab world (as our research context). section 2.1 presents educational data mining (EDM) and its associated techniques, as we will be using one technique (clustering) to analyse our data. This is followed by two chapters presenting two studies that inspired our research idea. The first study, presented in section 2.2, was published in 2013 by Kizilcec et al. [1], who analysed learner engagement with MOOCs on the Coursera platform. The second study, presented in section 2.3, was published in 2015 by Ferguson and Clow [2], who analysed learner engagement with MOOCs on the Futurelearn platform. In both chapters, we present a breakdown of their analyses and findings, which helped us in conducting our research. Chapter 3 shows the source of our obtained data, and explains the analytical methods used in our research. Chapter 4 shows the process of identifying the Arabic-speaking learners' profile. Chapter 5 shows the process of identifying the engagement types of the Arabic-speaking learners using Kizilcec et al.'s approach. Chapter 6 compares the Arabic-speaking learners with English-speaking learners from Kizilcec et al.'s study. Chapter 7 shows the process of identifying the engagement types of the Arabic-speaking learners using Ferguson and Clow's approach. Chapter 8 compares the Arabic-speaking learners with English-speaking learners from Ferguson and Clow's study. Finally, chapter 9 concludes this research with a summary of the key findings, research limitation, general discussion and suggestions for future work. Figure 0-1 illustrates the process of this research and some suggested areas for future work.

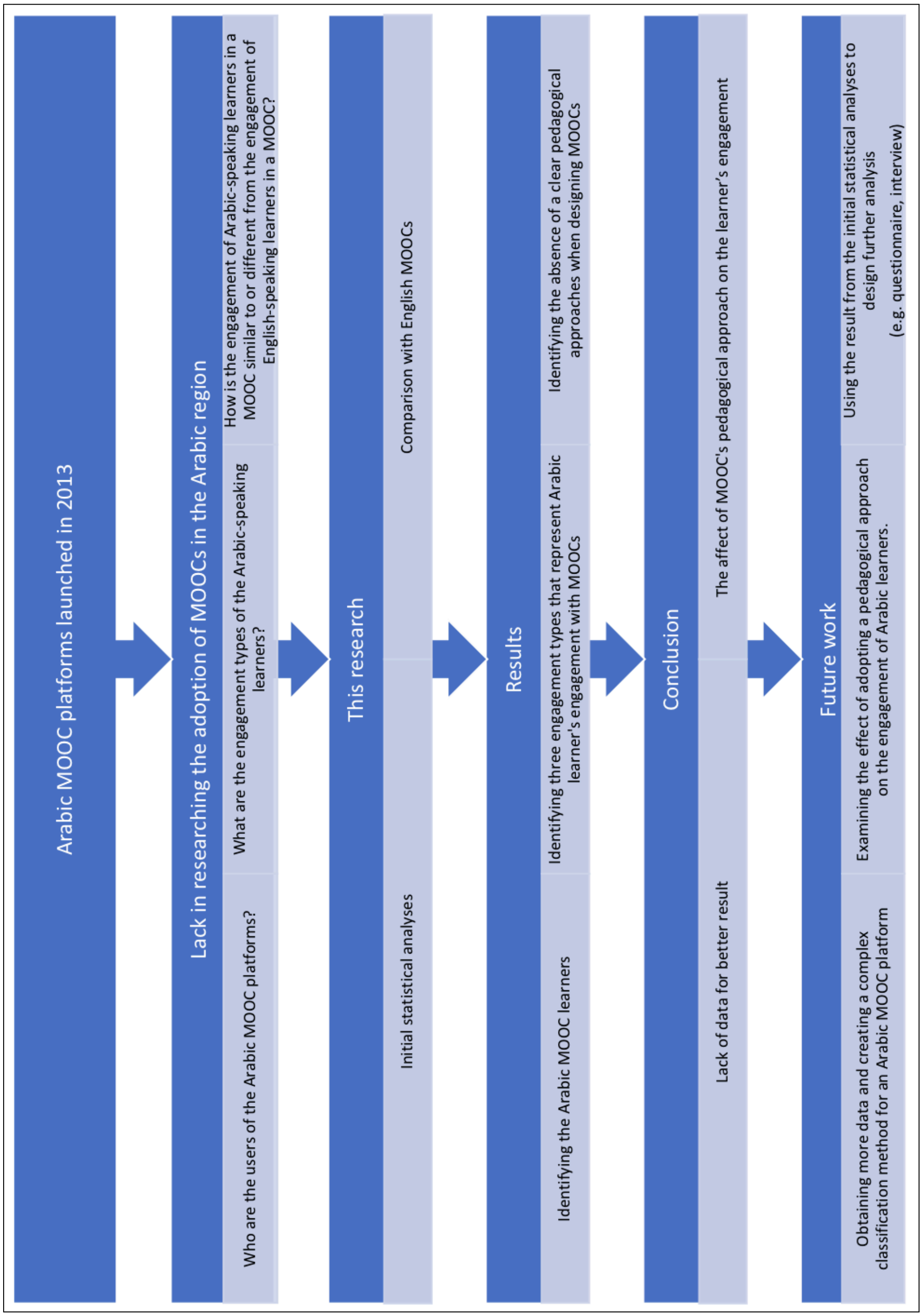


Figure 0-1: A flowchart of the research process

Chapter 1

Massive Open Online Courses (MOOCs)

As mentioned earlier, we studied the use of MOOCs in the Arab world. This chapter presents more information about the history of MOOCs and how it reached its current stage. Moreover, this chapter present some studies that investigated the use of MOOC in the Western world and the importance of such studies in enhancing the MOOC experience. We also pointed to the lack of such studies in the Arabic world.

1.1 MOOC Definition

MOOCs are defined as online courses that allow a large number of people to enrol with no qualification requirement or fees, and which are accessible by any user from any place with the aid of an internet connection [17]. The name was created by David Cormier to describe the first MOOC, Connectivism and Connective Knowledge, which was provided by George Siemens and Stephen Downes at the University of Manitoba in 2008 [7]. As Plourde [18] stated, "every letter is negotiable"; the first two words in the phrase (Massive Open Online Course) represent key features that distinguish MOOCs from other online courses:

- Massive: indicates the enormous number of participants that can be enrolled in a MOOC, unlike online courses in which the number of participants is limited [19].

- Open: addresses three aspects: (1) open access to all the course materials; (2) free registration to participate and learn, not to earn credit; (3) all work and materials are shared publicly for discussion [19].
- Online: addresses the method of delivering a MOOC, where the materials, assessments and discussions are delivered via the internet [19].
- Courses: indicates that it is a course in which curriculum materials are delivered to learners within a specific period of time [19].

1.2 MOOC Elements

MOOCs are offered by academic institutions or educational organisations [7] and presented on platforms such as Coursera, edX, Udemy, Futurelearn and Udacity. Each MOOC has a course description page, containing the course information, enrolment dates, deadlines and duration (usually from 6 to 14 weeks) [20]. Three main elements are found in most MOOCs: video lectures, assessments and discussion forums [21]. Video lectures are short (from 3 to 30 minutes) and provide learners the educational materials on a weekly bases [19]. Assessment in MOOCs has many forms, as assessing the progress of large numbers of learners is challenging [21]. Most MOOCs use automated assessment [19], either weekly quizzes or final exam [22]. Some MOOCs use peer assessment, where learners assess each other's work, or self-assessment [21].

During a MOOC, learners are encouraged to participate in discussions about the topics they are learning. MOOC platforms usually provide a discussion forum to facilitate communication between not only the learners but also the academic staff, who are usually required to be available for questions and feedback [20]. Completing a MOOC means successfully completing a specific percentage of the course that is decided on by the course provider. After completing a MOOC, some platforms allow learners to receive a certificate of completion for free (e.g., Udemy [23]). Other platforms, such as Coursera [24], edX [25] and Futurelearn [26], require additional fees to receive a certificate of completion.

1.3 MOOC History, Types and Pedagogy

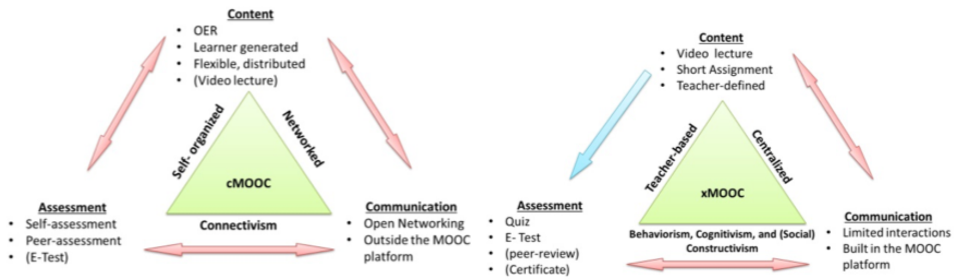
In 2008, George Siemens and Stephen Downes launched the first MOOC, entitled "Connectivism and Connective Knowledge (CCK08)", and attracted 2,300 learners [7, 27]. This MOOC was about connectivist pedagogy, a new learning theory advanced by George Siemens that emphasised the use of technologies and online connection in the educational process [28]. After CCK08, several MOOCs were launched, such as: Connectivism and Connective Knowledge (CCK09), Personal Learning Environments Networks and Knowledge (PLENK10), Connectivism and Connective Knowledge (CCK11) and Learning and Knowledge Analytics (LAK11) [29]. These MOOCs were termed Connectivist MOOCs (cMOOCs). In cMOOCs, the connections and discussions between participants is key for developing the learning process [30]. There is no formal assessment and no specific online platform used, instead, cMOOCs rely on Open Educational Resources (OER), webcasts, blogs and social media platforms [3, 30]. cMOOCs were designed using the connectivist principles of diversity, openness, interactivity and autonomy that were established by Siemens [28]. Stephen Downes explained the connectivist principles in cMOOCs as follows [30]:

- Diversity: in the tools used in the learning process, in learners' knowledge levels and in the content.
- Openness: in accessing the course materials and contents, in addition to open activities and assessment.
- Interactivity: in the collaboration and communication between learners to develop knowledge.
- Autonomy of learners: where learners decide on topics, contents and skills to learn, which eliminates the need for formal curriculum.

In 2011, a new MOOC on Artificial Intelligence was launched by Stanford University and successfully attracted 160,000 participants [29]. This was the first of the extended MOOCs (xMOOCs), which represent MOOCs that do not follow the connectivist pedagogy. Instead, xMOOCs follow a different pedagogy that was described by Siemens [19, p.7] as "teacher as expert and learner as knowledge consumer". The success of the Artificial Intelligence MOOC inspired its creators to start up a for-profit platform called Udacity to produce more MOOCs [29, 31]. In 2012, this drew the attention of educational institutions in the United States; Coursera was launched, one of the leading platforms for delivering MOOCs, after a partnership was formed with more than 30 educational institutions [32]. A month later, the EdX platform launched in cooperation with MIT and Harvard [3]. By December 2012, the Open University in the United Kingdom launched Futurelearn to provide MOOCs from leading universities in the UK [33], making the year of 2012 “The Year of the MOOC” [31].

Although xMOOCs differ in their pedagogical approach from the original MOOCs, cMOOCs, they both aim to create learning networks on a large scale; in this way, they are different from traditional learning in classrooms [3]. Yousef et al. [3] described the differences between the connectivist pedagogy in cMOOCs and the cognitive-behaviourist and social constructivist pedagogies in xMOOCs in Figure 1-1. They highlighted the follow key concepts [3]:

1. **MOOC organiser:** cMOOCs allow learners to self-organise their learning objectives. xMOOCs rely on the teachers to predefine the learning objectives.
2. **Content:** learners in cMOOCs use open educational recourses (OER) and share what they find useful and relatable to their course topic. xMOOCs use the teachers’ predefined materials as the main source for the course.
3. **Assessment:** cMOOCs self- or peer-assessments are informal and only used to improve understanding. xMOOCs peer- or E-assessment can be formal and result in obtaining a certificate.
4. **Communication:** learners in cMOOCs build their own networks outside the MOOC platform, using tools such as blogs, wikis and Facebook. xMOOCs’ learners communicate using a discussion forum that is provided on the MOOC platform.



(a) cMOOC key concepts

(b) xMOOC key concepts

Figure 1-1: Key concepts of cMOOCs and xMOOCs, from Yousef et al. [3]

Siemens [19] divided the pedagogical approach of the xMOOCs into two types: 1) cognitive-behaviourist pedagogy that is commonly used on platforms such as Coursera and edX [34]; 2) social-constructivist pedagogy that can be seen on the Futurelearn platform [2].

The cognitive-behaviourist pedagogy evolved from the concept of linking the occurrence of learning in humans with the adoption of new behaviours or changing existing behaviours [35, 36]. It calls for a scientific approach to be used for the development, application and assessment of learning [35]. The internet has helped to spread this pedagogy as a means of learning, where sharing content can be achieved in many ways, e.g., MOOCs [35, 37].

On the other hand, social-constructivist pedagogy supports group learning, where learners search for knowledge and share it [35, 37]. Educational tools and resources are used by learners as a shared means of learning and supporting one another [35, 36]. In social constructivism, knowledge is constructed through discussion, validation and application of knowledge in realistic situations [37]. Learners assess their learning when they develop the ability to self-evaluate [35]. The popularity of social constructivism has increased with the availability of communication technologies that mimic the traditional structures of a classroom [35, 37]. Table 1.1 shows how these three pedagogical approaches differ in knowledge construction and evaluation.

Table 1.1: Differences in knowledge construction and evaluation format between MOOC's pedagogical approaches

Pedagogical approach	knowledge construction tools	Evaluation
Connectivist	Open Educational Resources (OER), webcasts, blogs	Peer- or self-assessment Not formal
Cognitive-behaviourist	Videos and articles (predefined by instructor)	Peer- or E-assessment Can be formal and Result in obtaining certificate
Social-constructivist	Videos and articles (predefined by instructor) Supports group learning through discussion, validation and self-evaluate	Peer- or self-assessment Can be formal and Result in obtaining certificate

1.4 MOOCs in the Western World

The rapidly increasing number MOOCs has drawn the attention of higher education institutions [38] and motivated researchers to study MOOCs from different perspectives to enhance and improve the learning experience. The leading MOOC platforms in the Western world have been the subjects of these studies. Despite the increased popularity and the high enrolment in MOOCs, the majority of learners failed to stay engaged to the end of the course. This has made investigating the high dropout rate a priority for many studies. Adamopoulos [39] investigated the dropout rate through qualitative and quantitative analyses of data from multiple platforms. Adamopoulos identified several factors that have an effect on whether a learner completes a MOOC or not. Teachers, peer assessment and free textbooks were the main factors that had a positive effect on completing a MOOC. In contrast, the level of difficulty, workload, the duration of a course and the self-paced designed courses (that do not follow a specific timetable) had a negative effect on completing a MOOC [39].

Kizilcec et al. [1] investigation of the high dropout rate focused on one platform, Coursera. Kizilcec and his team [1] found that classifying learners into either completers or non-completers hides the reasons behind learners' choice to stop engaging with a MOOC. Therefore, their approach was to classify learners, using data analytics techniques, into groups based on their engagement with the course content [1]. After successfully identifying the engagement types of Coursera's learners, Ferguson and Clow [2] followed the same approach to identify the engagement types of Futurelearn's learners. In comparing their results with those of Kizilcec et al. [1], Ferguson and Clow [2] concluded that the platforms' pedagogical approaches affected the learners' engagement with the MOOCs.

Many studies were interested in researching the factors affecting learner engagement, especially learner completion. Methods ranged from using qualitative approaches, e.g., questionnaires and interviews, to quantitative approaches that analysed learners' actions on the platforms, i.e., video hits, quiz attempts or forum participation. Alraimi et al. [40] goal of identifying factors that enhance learners' motivation to continue using MOOCs led them to link MOOC completion to the reputation, openness and usefulness of the top university that provided the MOOC platforms. Gameel [41] concluded that usefulness, along with the interaction with the MOOC content, are the main factors in learner satisfaction.

Other studies focused on identifying factors that could help to predict learners' progress. Pursel et al. [42] found that completing a MOOC can be predicted by different indicators, such as: 1) high engagement with the course content; 2) previous educational knowledge; and 3) learner expectations from the MOOC. De Barba et al. [43], on the other hand, found that learner participation is the strongest predictor of performance. Moreover, they concluded that adding learner motivation to the analysis helped in better predicting the overall performance, making learners' participation and motivation key predictors of MOOC performance [43]. A similar result was found in Huang and Hew [44] work; they also identified a positive correlation between learner motivation and completion of the MOOC.

Despite the large number researches studying the Western MOOC platforms, there is no investigation on how the MOOC platform design is related to the teaching style that learners are familiar with in traditional classroom. In Western countries, such as the United States or United Kingdom, teaching style in traditional classrooms are based on interactive education that relies on instructor-student interactions [45]. Students in western countries are more encouraged to apply and practice various learning activities beside lectures including; discussions, note taking, case study analysis, field trips, and independent reading. In addition, exams/assignment methods are based on complicated practical problems [45]. Studying the relationship between the design of Western MOOC platforms and the teaching styles in classrooms might help explaining the learners interaction profile.

1.5 MOOCs in the Arabic World

The spread of MOOCs reached the Arabic-speaking region in 2013 when two Arabic language MOOCs were launched, Rwaq.org and Edraak.org. In September 2013, Rwaq launched as the first MOOC platform that designed and delivered its courses in the Arabic language [14, 15]. To provide top quality MOOCs, Rwaq partnered with major companies and agencies such as Microsoft, the Saudi Food and Drug Authority (SFDA) and the Arab League Educational, Cultural and Scientific Organisation (ALECSO). Rwaq now has more than 300 multidisciplinary courses and 34,844 daily visitor [46].

In November 2013, the Queen Rania Foundation for Education and Development (QRF) and the Open edX platform collaborated to launch the second Arabic-language MOOC platform, Edraak [47, 48]. They began by translating edX MOOCs to Arabic and providing them on the platform; today, Edraak is developing and producing its own MOOCs in partnership with top Arabic-language universities [14], such as University College London (UCL), American University of Beirut (AUB), American University in Cairo (AUC) and Princess Sumaya University for Technology (PSUT). In 2019, Edraak celebrated its fifth anniversary with more than two million learners and 110 courses [49], and it has 37,579 daily visitor [50].

After Rwaq and Edraak, several MOOC platforms were founded, each targeting a specific audience. In 2014, Nadrus.com was launched to provide both free MOOCs and paid MOOCs on multiple topics [14, 51]. Then, Doroob.sa was launched by the Human Resource Development Fund (HRDF) in Saudi Arabia following collaboration between edX and the Kingdom of Saudi Arabia's Ministry of Labour [52]. Although anyone can enrol in Doroob's MOOC, the purpose of this platform is to help unemployed Saudi citizens to gain the knowledge and skills required by many workplaces in Saudi Arabia. In addition, learners with completion certificates from Doroob are given priority in the hiring process at many national companies; however, although Doroob's MOOCs are free, requesting a certificate may be subject to some fees [53].

In 2015, the founders of Rwaq launched another platform called Maharah.net, or "Skill" in Arabic, that provides both free and paid MOOCs on multiple skills. The founders decided to develop this platform to separate the academic MOOCs they provide in Rwaq from skills-oriented MOOCs [54, 55]. Zadi.net is another MOOC platform that was launched in 2015. It is the first MOOC platform that specialises in Islamic and Sharia law [56]. The latest MOOC platform launched was in 2016, called Aanaab.com, and was developed following a collaboration between Rwaq and Emkan (an education development company). Aanaab is interested in providing academic MOOCs throughout the educational field. It targets educators, students and anyone with an interest in education in general [57, 58].

Rwaq, Edraak and Aanaab have start and end dates for each MOOC. They archive the MOOC when it is finished so that learners can access it to learn but without completing the assessment or requesting a completion certificate. The other Arabic-language platforms (Nadrus, Doroob, Maharah and Zadi) have a start date and sometimes an end date, but once a MOOC is released it stays on the platform and learners can enrol at any time in the future. These platforms focus on knowledge sharing more than course completion; therefore, only some MOOCs auto-assess the learners at the end of the MOOC and allow for certificates to be awarded. Table 1.2 summarises the Arabic-language MOOC platforms that were described in this section.

Table 1.2: Arabic MOOC Platforms

Platform	Founded	Number of MOOCs	Daily visitors*	Topics
Rwaq.org	2013	300	34,844	Multidisciplinary
Edraak.org	2013	110	37,579	Multidisciplinary
Nadrus.com	2014	153	354	Multidisciplinary
Doroob.sa	2014	114	695	Multidisciplinary
Maharah.net	2015	294	N/A	Multidisciplinary
Zadi.net	2015	139	3205	Islamic and Sharia law
Aanaab.com	2016	45	893	Education

N/A = Data not available *From CuteStat.com

MOOCs helped many societies, from spreading education on a massive scale at low cost [27, 59], to receiving high-quality education from top university without the constraints of time or place [60] and enabling communication and knowledge-sharing between different cultures [61]. The advantages of MOOCs have encouraged Arabic-language institutions that are interested in providing free education to launch Arabic MOOC platforms. However, compared to the developed countries, MOOCs, and open educational resources in general, are in their early stages [62, 63], and many attempts to develop open education failed due to the lack of research and limited vision [63]. As seen previously in section 1.4, researching MOOCs helped in the identification and prediction of useful information that can guide MOOC providers and developers to understand the learners' needs and improve the MOOC platforms accordingly. The Arab world lacks studies that analyse and accordingly improve Arabic MOOC platforms to suit the needs of its learners.

In the literature, we found only three studies that examined Arabic-language MOOCs. Adham [64] studied the use of avatars to represent female tutors in MOOCs, to examine the sociocultural barriers in gender-segregated societies. She found that using an avatar to represent a female tutor had a positive effect in overcoming cultural and social barriers in Saudi Arabia [64, 65]. Hakami [66] examined the factors affecting the continued use of Arabic-language MOOCs. She identified a set of factors that had a positive relationship on the continuance intention of learners in Arabic-language MOOCs. These factors were intrinsic motivations, perceived usefulness, Arabic language support, perceived ease of use and perceived reputation [66]. Finally, Almuhanna [67] investigated the perceptions and the impact of MOOCs on Saudi learners. She concluded that a MOOC's flexibility and the contribution to the development of educational cultures allowed it to be widely used, especially by women in Saudi Arabia. Moreover, she found that MOOCs helped to improve learners' knowledge and personalities, and to develop learners' educational and professional lives [67].

These three studies, [64], [66] and [67], used mixed methods of surveys, observation, interviews and focus groups. Moreover, the participants in these studies were Saudis and/or learners on the Rwaq platform. This shows the importance of having a more diverse study that examines Arabic-language MOOC platforms to benefit Arabic-speaking learners. The field of Arabic-language MOOCs is lacking a detailed analyses of the learners' interaction with the various parts of a MOOC, and not only focusing on completion. More studies should aim to investigate Arabic learners' objectives in enrolling in a MOOC, to help the MOOC's providers and the platform's designers to understand and meet the needs of the Arabic-language learners.

In addition, exploring the link between the design of Arabic MOOC platform and the teaching style in traditional classrooms might provide a better understanding of learners interaction with MOOC contents. In Arabic classrooms, the education system is known to be an examination-oriented system that evaluate the passive absorption of knowledge [45]. The teaching style followed in educational institutions in the Middle East depends mainly on lectures, rote learning, and teacher dictation [68, 69]. The Teaching process is based on illustrating concepts and reading from textbooks [45, 70]. Students are expected to follow teachers instructions as an absolute authority. Furthermore, students are not encouraged to learn about issues management unless they directly affect their curriculum [71], nor to engage in an interactive or group activities [70]. Exams are the main assessment method for students and relies on memorizing facts and not on applying concepts. Exams does not normally contain questions require students to apply what they have learned outside the class [72]. In accordance to these learning and assessment activities, students in the Middle East counties prefer to be told what to study, which materials to read and what to pay attention to, and favor courses that offer short summarized sources as reading material [70], and reading materials that contain simple and clear information that can be memorized easily [45, 69].

Chapter 2

Review of User Engagement with English MOOC

2.1 Educational Data Mining

As mentioned earlier, in the objective section on page 3, we used K-means clustering algorithm to analyse learner data from an Arabic-language MOOC platform. Therefore, in this chapter we briefly review the field of Educational Data Mining (EDM), where clustering is one of its methods. In addition, we explain how K-means clustering works.

2.1.1 The Beginning of EDM

EDM, as explained by the Journal of Educational Data Mining, is "an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings in which they learn" [73, para. 2]. Siemens and Baker [74] linked the emergence of EDM to two causes: 1) the emergence of "big data" after employing new educational media [75]; and 2) the availability of very large data sets generated from educational software and online learning [76]. In 2005, EDM started as a workshop series that was turned into an annual conference in 2008, which led to issuing the first publication of its journal in 2009 [77]. Romero et al. [78] summarised the educational applications of EDM as follows:

- Communicating to stakeholders: helping educators to analyse learner engagement with the course.
- Maintaining and improving courses: helping educators to determine ways to improve the course using the learners' usage information.
- Generating recommendations: helping learners find the best content that is appropriate for them at a specific time.
- Predicting learner grades and learning outcomes: based on the data of learners' interaction with the course content.
- Learner modelling: helping to create models from learners' usage information, to identify satisfaction, motivation and learning progress.

2.1.2 EDM Methods

EDM includes a wide range of methods. The four most frequently used methods are: Prediction, Relationship Mining, Structure Discovery and Discovery with Models [77]. In this section, we explain these four methods and the purpose for using each.

Prediction

Prediction refers to developing a model that resolves one variable of the data by combining other variables [77]. Prediction is commonly used to predict values in specific contexts, e.g., learners' future outcomes [79]. There are three types of prediction:

- Classification, where the predicted variable can be binary or categorical [77]. Common classification measurements are [80], Precision and Recall [81].
- Regression, where the predicted variable is continuous [77]. Common regression methods in EDM are linear regression and regression trees [77].
- Latent Knowledge Estimation, where "a student's knowledge of specific skills and concepts is assessed by their patterns of correctness on those skills (and occasionally other information as well)" [77, p. 64].

Relationship Mining

Relationship mining is about searching big data sets to find relationships between variables, whether they are one-to-one variable or one-to-many variables [77]. There are four types of relationship mining:

- Association rule mining, where the relation between variables follows the rule of "if-then", e.g., IF a learner wants to learn, THEN the learner asks for help [77]. Application of association rule mining can be found in [82] and [83].
- Correlation mining, where the relation between variables is either a positive or a negative linear correlation [77]. Correlation mining has been used to link student attitudes to help-seeking behaviours [84], and to study the design of intelligent tutoring systems and gaming the system [85].
- Sequential pattern mining, where the relation between variables shows the temporal associations between events [77]. Sequential pattern mining was used by Perera et al. [86] to study learner behaviours in collaboration work, and by Shanabrook et al. [87] to study the patterns in help-seeking behaviour.
- Causal data mining, where the relation between variables shows whether an event is caused by another event [88]. Causal data mining has been used in EDM in predicting factors leading to learners' poor performance [89], and in studying the impact of gender and attitudes on intelligent tutor behaviours [90].

Structure Discovery

In structure discovery, the aim is to develop algorithms that discover a structure in the data without help from prior knowledge or a basic truth [77]. Three structure discovery algorithms are used in EDM:

- Clustering, where the aim is to divide the data into a set of clusters. The data in each cluster share a common feature that does not exist in other clusters [91]. There are two types of clustering: hierarchical (e.g., Agglomerative and K-D trees), and non-hierarchical (e.g., K-means and Spectral clustering) [77].

- Factor analysis, where the aim is to divide the data into a set of factors [92]. Factor analysis is used to validate or determine scales [77].

- Domain structure discovery is about "finding which items map to specific skills across students" [77, p.68]. Domain structure discovery uses automated algorithms (see Thai-Nghe et al. [93]), or allows for human judgment (see Cen et al. [94]).

Discovery with Models

Discovery with models is about developing a model using prediction or clustering techniques, and then using this model as a part of another analytics technique such as prediction or relationship mining [77]. Common uses of the discovery with models method in EDM include using predictions of a primary model as a predictor in a new model (see Baker et al. [95]).

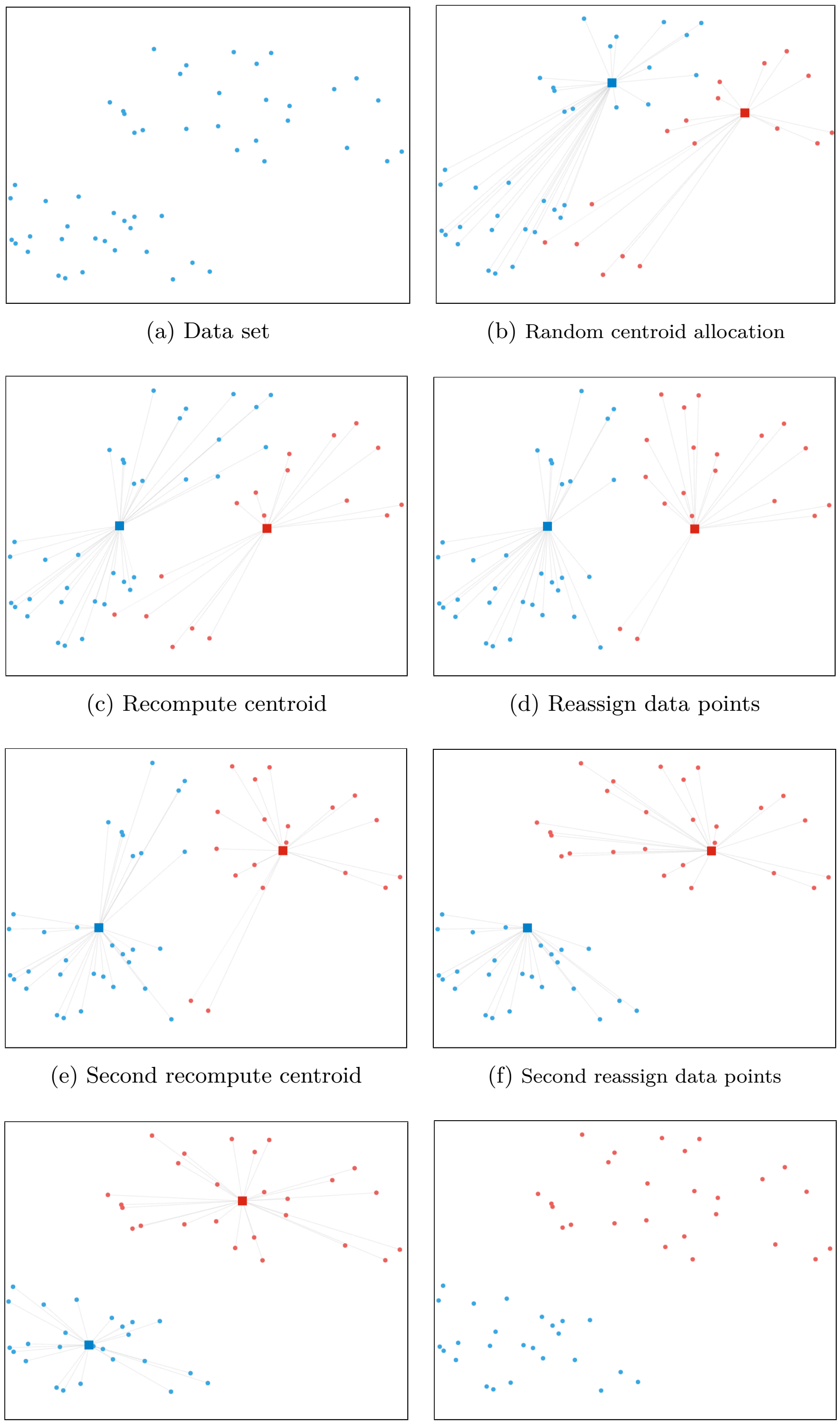
2.1.3 K-means Clustering

This study employed the K-means clustering algorithm to classify Arabic-language MOOC learners. Therefore, this section explains how this algorithm works. As mentioned above, clustering is one of the Structure Discovery methods and is categorised into hierarchical and non-hierarchical algorithms. K-means is a non-hierarchical clustering algorithm that was first used by MacQueen et al. [96] in 1967. MacQueen et al. explain the K-means clustering algorithm as a method for classification of multivariate observations into a K number of clusters [96].

This algorithm works by first identifying the number of clusters K in which we want our data set S to be clustered. Then the algorithm places K number of centroids into the data at random locations. Next, each data point in the data set S is assigned to the nearest centroid by computing the distance between the data point and all the centroids. This computation is performed using the Euclidean norm (L_2 norm), which measure the distance between two points in euclidean space (x,y). The L_2 norm, see equation 2.1, calculates the square root of the sum of the squared differences of the two points' coordinates. After assigning all the data points in S to the K clusters, the mean of each cluster is computed and identified as the new centroid of its cluster. Then the algorithm repeats the last two steps of computing the distance between all data points and the new centroid and assigning the data point to the nearest, and recomputing the means of each cluster and identifying it as the new clusters' centroid. The repetition continues until convergence, when there are no more changes in the assignment of the data points to the clusters and the centroids do not change [96]. The following example shows clustering a data set using K-means with K = 2.

$$d(A, B) = \sqrt{\sum_{i=1}^{n} (A_i - B_i)^2} \tag{2.1}$$

Example. *When clustering the data set in Figure 2-1a using K-means, first we identify the number of clusters K. For this example, we chose $K = 2$. Then, the algorithm places 2 centroids into the data at random locations. Each data point in the data set is assigned to the nearest centroid by computing the distance between the data points and the two centroids (see Figure 2-1b). Then, the mean of each cluster is computed and identified as the new centroid of its cluster (see Figure 2-1c). The algorithm then continues the steps of recomputing the distance between all data points and the new centroid and assigning each data point to the nearest centroid, and recomputing the means of each cluster and identifying it as the new clusters' centroid (see Figures 2-1d, 2-1e, 2-1f and 2-1g). After the third recomputation of the centroid, convergence is achieved and we have our final two clusters (see Figure 2-1h).*



(a) Data set

(b) Random centroid allocation

(c) Recompute centroid

(d) Reassign data points

(e) Second recompute centroid

(f) Second reassign data points

(g) Third recompute centroid

(h) Final clusters

Figure 2-1: The process of K-means clustering algorithm with K=2

2.2 User Engagement with MOOCs in Coursera

In chapter 1, we identified MOOC and examined its design and objectives to aid our understanding of how learners engage with it. We then explained the field of Educational Data Mining (EDM) in section 2.1. We looked at EDM methods, especially the K-means clustering algorithm; this algorithm will be further discussed in this and subsequent chapters. In this chapter, we will examine the fundamental work by Kizilcec et al. [1] in the area of learner engagement with MOOCs. We will present a summary and analysis of their work in analysing engagement patterns of MOOC learners in the Coursera platform (a platform based in the United States).

Kizilcec's team addresses the issue of classifying MOOC learners into two categories, completers and noncompleters, or, as Kizilcec et al. describe them, "those who pass the class by adhering to the instructor's expectations throughout the course, and everyone else" [1, p. 170]. Therefore, the focus of their analysis was to identify learners' engagement and disengagement with MOOCs. Their method included a mechanism for classifying learners' interactions with MOOCs, and a clustering algorithm for identifying engagement types. This chapter examines the possibility and the process of applying their techniques to learners' data on the Edraak MOOC platform. The chapter will conclude with some recommendations as to how we can adopt Kizilcec et al.'s methodology and apply it to Edraak's data.

2.2.1 Kizilcec et al.'s Data Source

Three computer science MOOCs from the Coursera platform were chosen by Kizilcec's team [1] for their study. Each MOOC covered the contents for a specific level of education: Computer Science 101 for the high school level (HS); Algorithms: Design and Analysis for the undergraduate level (UG); and Probabilistic Graphical Models for the graduate level (GS); see Table 2.1. The platform, Coursera, and the three courses were designed and delivered in the English language. Each of the three MOOCs ran for nine weeks, and had one assessment at the end of each week. Learners could interact with the course contents at any time and could also submit the assessment after the deadline.

Table 2.1: Summary of Kizilcec et al.'s data sample

	MOOC 1 (HS)	MOOC 2 (UG)	MOOC 3 (GS)
Subject	Computer Science	Computer Science	Computer Science
Level	High school	Undergraduate	Graduate
Duration	9 weeks	9 weeks	9 weeks
Number of Assessments	9	9	9
Number of Learners	46096	26887	21108

The majority of learners in the three MOOCs were located in the United States; the next two largest groups were from India and Russia, respectively. Male learners dominated all three courses, with 64%, 54% and 70% for high school, undergraduate and graduate, respectively. Kizilcec's team [1] also looked at the Human Development Index (HDI) of the learners' countries, based on the United Nations Development Programme report of 2011 [97].

Definition. *The Human Development Index (HDI) is an indicator used by the United Nations Development Programme to rank countries based on their health, knowledge and standard of living, into four levels of development (Very High, High, Medium and Low) [98]. Appendix A explains more about HDI and how it is calculated.*

In all three courses, the majority of learners were from countries having a Very High HDI level, with 69%, 54% and 70% for high school, undergraduate and graduate MOOCs, respectively, and very few learners from Low HDI countries, with 3%, 3% and 1% for high school, undergraduate and graduate MOOCs, respectively.

2.2.2 Kizilcec et al.'s Methodology

The analysis by Kizilcec's team had two stages [1]: identifying an engagement description for each learner, and applying a clustering algorithm to identify the engagement types. The following terms are used to discuss the analysis by Kizilcec's team:

- Course contents: refers to the video lectures and online assessment.

- Interaction with course contents: watching some or all video lectures and/or solving assessments.
- Assessment period: the time between assessments, i.e., weekly.

2.2.2.1 Kizilcec et al.'s Engagement Description

After each assessment period in a MOOC, an engagement description for each learner was computed based on the learner's interaction with course contents. Kizilcec et al. [1] focused on two content types: watching video lectures and solving assessments. They believed these two content types would be found in any MOOC, regardless of the platform design or pedagogical approaches followed. At the end of each week, learners were assigned one of the following letters, which reflected their interaction with the course contents that week [1]:

- When learners completed the assignment on time, they were "on track" and given the letter "T"
- When learners completed the assignment late, they were "behind" and given the letter "B"
- When learners did not complete the assignment but watched a video, they were "auditing" and given the letter "A"
- When learners had no participation at all, they were "out" and given the letter "O"

MOOCs used in Kizilcec et al.'s study [1] ran for nine weeks, which means each engagement description comprises nine letters. For example, if a learner watched video lectures and completed the assignments on time for the first four weeks, missed the fifth week, submitted the assignments for week six and was late for week seven, then watched video lectures only without submitting the assignments for weeks eight and nine, then the engagement description for this learner would be [T,T,T,T,O,B,B,A,A].

Even though Kizilcec's team [1] focused on two types of content, videos and assessment, we believe they did not consider all possible interactions with the course content.

For ‘late submission’, Kizilcec’s team assigned the letter ‘B’, and for ‘on-time submission’, they assigned the letter ‘T’. In neither case did they include the interaction with video lectures. Therefore, whether or not a learner watched a video, the learner would receive a ‘B’when submitting the assessment late, and the same for the letter ‘T’. We believe that with two types of course content, we can identify six interactions:

- Learners who watched video lectures and complete the assignment on time.
- Learners who did not watch video lectures but complete the assignment on time.
- Learners who watched video lectures and complete the assignment late.
- Learners who did not watch video lectures but complete the assignment late.
- Learners who watched video lectures but did not complete the assignment.
- Learners who did not watch video lectures and did not complete the assignment.

2.2.2.2 Kizilcec et al.’s Calculated Similarities

After forming an engagement description for each learner, Kizilcec’s team [1] assigned a number to each letter, as so: T = 3, B = 2, A = 1, O = 0. This converted the engagement description example of [T,T,T,T,O,B,B,A,A] to [3,3,3,3,0,2,2,1,1,]. While this was done to allow Kizilcec’s team to calculate the similarity between learners’ engagement descriptions, it makes the letters assigned at the beginning of the analysis useless. This can be avoided by assigning numbers directly to the learners’ interaction with course contents.

These numbers (T = 3, B = 2, A = 1, O = 0) indicate that Kizilcec’s team [1] assigned greater value to solving assessments than to watching video lectures. There is no justification for this from Kizilcec’s team; however, we believe it matches the traditional learning in a classroom, whereby passing the exam is more important than attending all the lectures.

To calculate the similarity between learners’ engagement descriptions, Kizilcec et al. [1] used the L_1 norm instead of the L_2 norm. See section 2.1.3 for more information about the use of L_2 norm in K-means clustering.

Drawbacks of the L_1 norm

Definition. *The L_1 norm is a geometric calculation of the distance between two points. It is reached by calculating the sum of the absolute differences of the two points' coordinates [99][100].*

To understand the L_1 norm equation, we will calculate the similarity between two examples of learners' engagement descriptions.

Example. *A is the first learner, with engagement description = [3,2,3,3,2,0,0,1,0]. B is the second learner, with engagement description = [1,2,2,2,0,1,2,2,0]. The L_1 norm equation is:*

$$d(A, B) = \sum_{i=1}^{n} |A_i - B_i|$$

where i is the week's number, n is the number of weeks in the MOOC, A_i is learner A's interaction with content per week, B_i is learner B's interaction with content per week and $d(A, B)$ is the distance between A and B. Applying this equation to learners A and B to calculate the similarity between them, we get the following:

$$d(A, B) = |3-1| + |2-2| + |3-2| + |3-2| + |2-0| + |0-1| + |0-2| + |1-2| + |0-0|$$
$$d(A, B) = |2| + |0| + |1| + |1| + |2| + |-1| + |-2| + |-1| + |0|$$
$$d(A, B) = 2 + 0 + 1 + 1 + 2 + 1 + 2 + 1 + 0 = 10$$

The step of applying the L_1 norm equation raises a concern for us related to the loss of the detailed weekly engagement, where learners with different engagement descriptions may have the same distance after the L_1 norm equation is applied. For example, if learner C's weekly engagement description is [0,1,1,3,1,1,1,2,0], then the similarity/distance between learners A and C (see the calculation below) equals the similarity/distance between learners A and B, even when each learner has a different engagement description representing different interactions with the course contents.

$$d(A, C) = |3-0| + |2-1| + |3-1| + |3-3| + |2-1| + |0-1| + |0-1| + |1-2| + |0-0|$$
$$d(A, C) = |3| + |1| + |2| + |0| + |1| + |-1| + |-1| + |-1| + |0|$$

$d(A, C) = 3 + 1 + 2 + 0 + 1 + 1 + 1 + 1 + 0 = 10$

When we asked them about their scoring system, Dr R Kizilcec (personal communication, October 9, 2018) responded with, ‘We choose the L_1 norm because it doesn’t spread out the distances too much. Why would being On Track be 9 times better than Out, but only 4 times better than Auditing? Distances of 3 and 2 seemed a bit more reasonable to us’.

2.2.2.3 Kizilcec et al.’s Clustering Algorithm

The second stage of Kizilcec et al.’s analysis was to apply a clustering algorithm. They chose a K-means algorithm based on previous uses of the method on actual courses (not online) in other studies [1]. One study used K-means clustering to classify community college students into six types, based on their enrolment patterns [101]. Another study used K-means clustering to classify community college students into fifteen types, based on their engagement, which was measured by a survey [102].

K-means clustering is one of the clustering algorithms used in EDM. In one study, it was applied to a data set to classify it into K number of clusters [96]. Each cluster was a group of points from the data set that shared one or more characteristics. When the K-means was run with K=3, for example, the algorithm generated 3 random points in the data set called centroids. The algorithm then calculated the distance between each data point and the three centroids. It assigned each data point to the nearest centroid, forming three clusters. The algorithm then reallocated the three centroids by calculating the mean of each cluster and recalculating the distance between each data point and the three new centroids. This process of reallocating the centroids and recalculating the distance was repeated until convergence, which means there were no further changes in the assignment of the data points to the clusters and the centroids did not change. Section 2.1.3 explains the K-means clustering algorithm in detail.

Kizilcec and his team ran the K-means clustering algorithm with K = 4, and repeated the clustering one hundred times to select the highest likely result [1]. The algorithm outputted four clusters, where each cluster comprised learners who shared the same engagement description. We developed the example below to provide a better understanding of the purpose of running K-means many times before selecting the highest likely result.

Example. *Let us assume a data set of some points on a two-dimensional space (X, Y). In Figure 2-2, we present two possible options, A and B, for the random allocation of centroids. After applying K-means to the same data set with the same value of $K = 3$, the final three clusters were different. Therefore, repeating K-means numerous times is important as it helps the algorithm to select the highest likely result.*

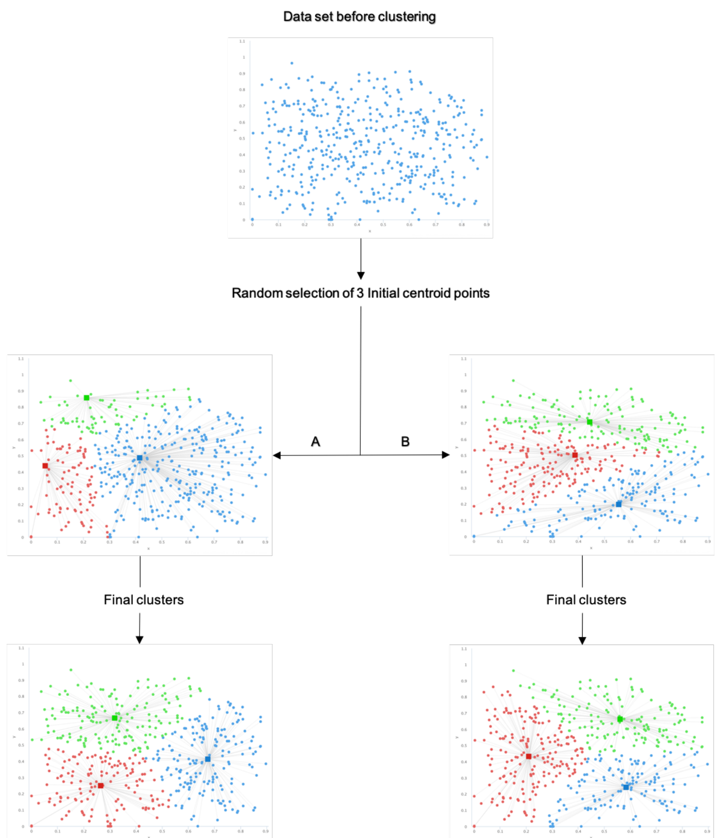


Figure 2-2: Two options for the initial centroid allocation in a K-means clustering algorithm with K=3

The issue of using K-means in Kizilcec et al.'s study [1] is that they used the L_1 norm equation. This makes their data one-dimensional, meaning that the data set becomes linear and cannot be represented in a Euclidean space. Although K-means could be applied to a one-dimensional data set, it is not the best algorithm for clustering such data [103][104].

On the other hand, the literature has suggested different types of classification for one-dimensional data. One example is the Kernel Density Estimation (KDE) method, which estimates the probability density function of random data points [105] and uses it to split the data set at the local minima points [106]. The Jenks natural breaks optimization is another classification method, which determines the optimal classification of data points [107]. Finally, Wang and Song [108] developed a new algorithm, an R package called Ckmeans.1d.dp, which they claimed is optimal for one-dimensional clustering. We will not further describe one-dimensional clustering, as the main issue here is the unconvincing use of the L_1 norm equation in the first place, since it meant the analysis lost some meaningful data.

2.2.3 Kizilcec et al.'s Findings

The result of Kizilcec et al.'s one-dimensional clustering [1] was four clusters that represented four types of engagements in the three MOOCs from the Coursera platform. Learners from three engagement types did not complete the MOOCs, while learners from the fourth engagement type did. The four engagement types (see Table 2.2) are as follows [1]:

1. Sampling: The majority of learners from the three MOOCs fell into this group, with an average of 64.3%. Learners watched only one or two video lectures at the beginning of the course before dropping out.

2. Disengaging: With an average of 15.3%, learners completed the assessments at the beginning, then gradually dropped out within the first third of the course.

3. Auditing: With an average of 7%, these learners engaged with the course for most of its duration only by watching video lectures and did not do the assessments.

4. Completing: With an average of only 13.3% of the three MOOCs, this is the only type where learners engaged fully by watching the majority of the video lectures and completing the assessments.

Table 2.2: Engagement types of MOOCs on the Coursera platform [1]

Course	Completing	Auditing	Disengaging	Sampling
MOOC 1	27%	6%	28%	39%
MOOC 2	8%	6%	12%	74%
MOOC 3	5%	9%	6%	80%
Average	13.3%	7%	15.3%	64.3%

Kizilcec et al. [1] ran the following three tests to evaluate their clusters and to ensure their methodology:

- **Robust clusters:** Their first test was to ensure that the clusters they had were strong enough to face minor changes in their methodology. They tried two minor changes, one related to the type of interaction measured, and the other related to the chosen value of K. The first change related to the first stage of the analysis, in which they assigned one of the following interactions to each learner: On track, Behind, Auditing, Out. They later chose to include a new interaction type they called Assessment pass and exclude Behind. The results showed similar clusters with a 95% overlap with the original clusters [1].

  The second change was to try a different number of clusters (a different K value). When they did so, they noticed that increasing the value of K over 4 only divided the original clusters into sub-clusters [1]. They gave an example of K = 5; when they ran this clustering on the undergraduate-level MOOC, they found that the original Sampling cluster was divided into two clusters, one for the learners who sampled during the first weeks of the MOOC, and the other for learners who sampled during the last weeks of the MOOC. This result allowed them to claim their four clusters represented the high-level engagement types of their MOOCs [1].

- **Goodness of fit:** Their second test was to check the goodness of fit of the data in their clusters [1]. This refers to how well each data point fit a set of data within the same cluster. To check this, they used Rousseeuw's Silhouettes validation test [109]. The Silhouettes validation test works by measuring a score between 0 and 1 for each data point. The closer the score is to 1, the closer the data point fits to the middle of the cluster. A score is measured as follows:

Example 1. *Let us assume two clusters, C1 and C2, where each has a set of data points (Figure 2-3). For point X in C1, the Silhouettes validation test measures the average distance between X and all points in C1; this is called distance A. It then measures the average distance between X and all points in C2; this is called distance B. The score for point X is measured by calculating the difference between B and A and dividing the result by the maximum of A and B.*

$$X = \frac{B - A}{MAX(A, B)}$$

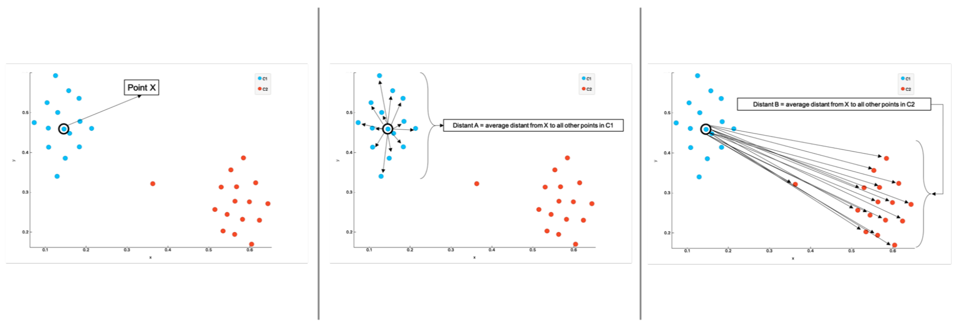


Figure 2-3: Silhouettes validation test

The Silhouettes validation test calculates the scores of all points and finds the goodness of fit for all points to their cluster. Figure 2-4 shows the points' scores. We can see that the blue point in the middle, which belongs to cluster 2, shows a low score. This means it does not have a good fit in C2. We can also see that the bottom point in C1 and the top point in C2 have lower scores than the rest of the points; this indicates they are the farthest from the centre of their clusters.

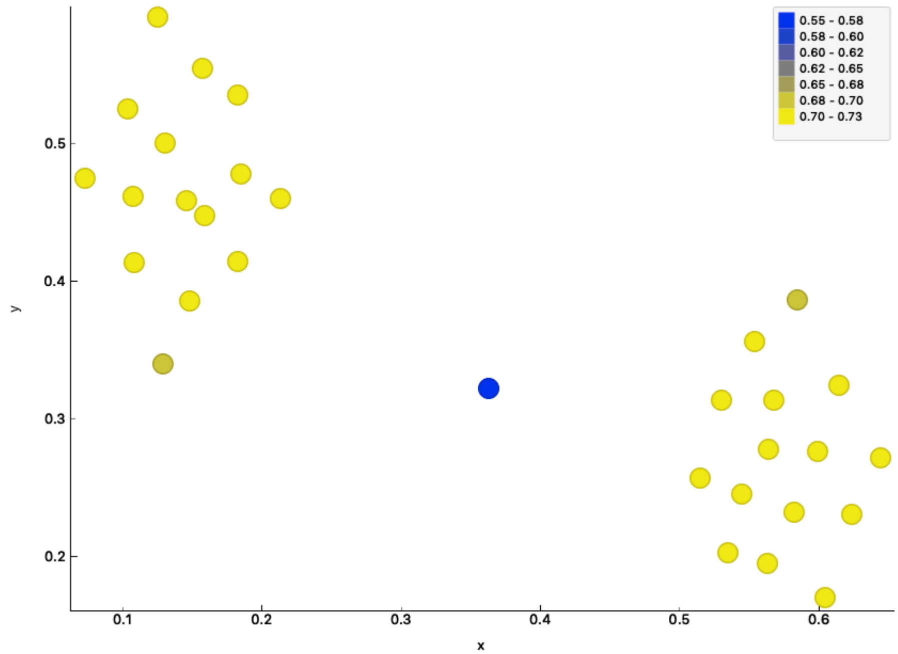


Figure 2-4: Data score in Silhouettes validation test

- **Educational sense:** The third and final test done by Kizilcec's team to evaluate their clustering methodology was to ensure the clusters had a pedagogical existence [1]. Kizilcec's team used a common sense test and asked the question, 'Is it plausible to imagine a posteriori that students would interact in an educational platform in these high-level ways?'[1, p.173]. They believed their engagement types (Completing, Auditing, Disengaging and Sampling) were meaningful, which was linked to the use of small, special labels (On track, Behind, Auditing, Out) [1].

2.2.4 Recommendations

We conclude this chapter with some recommendations that might provide a better analysis of learners' engagement with MOOCs. The following points were suggested by Kizilcec et al. [1] for future research, to further understand learner engagement:

1. **Scoring system:** Kizilcec's team believed that adding more course features, e.g., quiz and assessment marks, would allow the clustering algorithm to detect more engagement types.

2. **Assessment period:** They claimed that narrowing the assessment period, which was one week, to a smaller window (e.g., one hour) would allow them to investigate learners' engagement on a work session basis rather than a week. However, they had some concerns that adding complex features to the analysis might result in missing the big picture.

3. **Compare with other MOOCs:** Kizilcec's team also recommended future researchers use their methodology and compare with other MOOCs. This, they believed, would help to identify changes in learner engagement that are caused by differences in pedagogy approaches or learner experiences with MOOCs over time.

4. **Learners' experiences:** Kizilcec's team predicted that learner engagement with the same MOOCs would change in the future, once learners became familiar with the course design.

On the other hand, we believe that in order to adopt Kizilcec et al methodology and apply it on our Edraak data, we should consider the following:

1. **Same MOOCs levels:** Since we were using more than one MOOC, the levels in which the course content was selected had to be the same. Having different levels (high school, undergraduate, etc.) would prevent us from knowing the learners' specific requirements. Alternatively, a qualitative analysis would be needed to address each level of learners.

2. **Scoring system:** When assigning scores to the learners' interaction with the course contents, we had to consider all possible interactions. In addition, we had to justify why some interaction types were considered to have more value than others, e.g., solving assessments had more value than watching video lectures.

3. **Avoid using the L_1 norm equation:** It is not ideal for K-means clustering, which is a multidimensional algorithm. In addition, the L_1 norm discards very useful information about learner interactions and shortens the engagement description to a singular number.

4. **Clustering algorithms:** K-means is a good classification algorithm when used with multidimensional data. Therefore, using it with the full engagement description would provide a more detailed insight into learner engagement with MOOCs.

However, as done by Kizilcec et al. [1], we needed to run some validation tests to insure that the value of K (number of clusters) was optimal. Alternatively, hierarchical clustering was also a good classification algorithm that would not require specifying the number of clusters in advance.

2.3 User Engagement with MOOCs in Futurelearn

In section 2.2, we explained the work of Kizilcec et al. [1] in analysing MOOCs on the Coursera platform. In this chapter, we present how another research team adopted Kizilcec et al.'s methodology and applied it to a different platform. Ferguson and Clow [2] studied learner engagement with MOOCs on Futurelearn, a platform based in the United Kingdom. They followed the same approach of identifying engagement descriptions, then clustering the data using a K-means algorithm. However, they found that Kizilcec et al.'s approach did not suit the design of the Futurelearn platform. Moreover, they addressed the issue of using K-means as a one-dimensional algorithm. Ferguson and Clow [2] applied some changes to Kizilcec et al.'s approach [1], aiming to show that learner engagement with MOOCs is influenced by the pedagogical approach used to design the MOOCs on the Futurelearn platform.

In this chapter, we review Ferguson and Clow's work [2] and compare it with Kizilcec et al.'s work [1]. This allowed us to improve our understanding of how English-language MOOC platforms are designed and the best way to analyse learner data. It also guided us to adopt the approach that would fit the design of the Edraak MOOCs and to analyse Arabic-language learner engagement. Similar to the previous chapter, we conclude with some recommendations on how we can adopt Kizilcec et al.'s methodology, considering the modifications made by Ferguson and Clow [2], and how to apply it to Edraak's data.

2.3.1 Ferguson and Clow's Data Source

Four MOOCs from the Futurelearn platform were chosen by Ferguson and Clow [2] for their study. Unlike those chosen for Kizilcec et al. [1] study, the four MOOCs were all designed for the beginner's level and required no previous experience; however, each represented a different subject area. MOOC 1, Physical Sciences, ran for eight weeks and included one assessment per week. MOOC 2, Life Sciences, ran for six weeks and included one assessment per week. MOOC 3, Art, ran for eight weeks but included only three assessments during the course. MOOC 4, Business, was similar to MOOC 1 in that it ran for eight weeks and included one assessment per week [2]; see Table 2.3. The Futurelearn platform and the four courses were designed and delivered in the English language. Learners could interact with the course contents at any time. Moreover, there was no deadline for the assessments, so learners could submit at any time before the end of the course. The majority of learners in the four MOOCs were located in the United Kingdom, as the Futurelearn platform was newly launched. Female learners dominated in these courses. with 48%, 61%, 67% and 65% for MOOCs 1 to 4, respectively.

Table 2.3: Ferguson and Clow's data sample

	MOOC 1	MOOC 2	MOOC 3	MOOC 4
Subject	Physical Sciences	Life Sciences	Art	Business
Level	Beginner	Beginner	Beginner	Beginner
Duration	8 weeks	6 weeks	8 weeks	8 weeks
Number of Assessments	8	6	3	8
Number of Learners	5069	3238	16118	9778

2.3.2 Ferguson and Clow's Methodology

Ferguson and Clow [2] attempted to adopt Kizilcec et al.'s analysis [1] without any change at first; however, the result was not clear. In this section, we summarise Ferguson and Clow's first attempt and describe the issues they believed affected their result. We then explain the changes made to the methodology that makes it applicable to the Futurelearn platform. Some terms that will be used to discuss Ferguson and Clow's analysis include:

- Course contents: refers to the video lectures and online assessment.
- Interaction with course contents: watching some or all video lectures and/or solving assessments.
- Assessment period: the time between assessments, i.e. weekly.

Ferguson and Clow's Initial Attempts

Ferguson and Clow [2] began by following the exact approach of Kizilcec et al. [1], by assigning the same labels (On track, Behind, Auditing and Out) to their learners at the end of each assessment period (i.e., weekly), based on the learners' interaction with two types of course content (watching video lectures and completing assessments).

They then gave each label the same weight suggested by Kizilcec's team (On track = 3, Behind = 2, Auditing = 1 and Out = 0) to form an engagement description for each learner. Next, like Kizilcec's team, Ferguson and Clow [2] applied the L_1 norm equation and used a K-means clustering algorithm with K = 4 in an attempt to replicate Kizilcec et al.'s result. The resulting engagement types from Futurelearn MOOCs showed that two of the four clusters were similar to the Completing and Sampling clusters from Kizilcec et al.'s study; however, the other two clusters showed significant overlap.

Following Kizilcec et al.'s approach to validate their clustering, Ferguson and Clow [2] measured the four clusters' Silhouette scores and found the average score to be 0.67; this did not indicate a good fit (for more on the Silhouettes validation test, see example 1 on page 34). Therefore, they decided to repeat the clustering with different K values. They repeated the clustering with K = 3 to 8, and found that the Silhouette scores decreased even more. Ferguson and Clow [2] then addressed the issue of using K-means as a one-dimensional algorithm, where this dimension was the engagement description summed in a singular variable. The issue of one-dimensional K-means clustering was discussed in section 2.2.2.3. Ferguson and Clow [2] believed that the one-dimensional approach missed useful data about learner engagement with MOOCs. Again, Ferguson and Clow [2] repeated their clustering with K = 4; this time, however, they skipped the step of applying the L_1 norm equation and used the full engagement description to apply the K-means as a multidimensional algorithm (six- and eight-dimensional for the six-week MOOC and the eight-week MOOC, respectively). However, the result showed only two clear clusters, Completing and Sampling; the other two clusters were unclear. They then repeated the multidimensional approach with K = 3 to 8, but again the Silhouette scores decreased even more than with the one-dimensional approach.

2.3.2.1 Ferguson and Clow's Engagement Description

After achieving mixed results from their initial attempts, Ferguson and Clow [2] decided to make some changes to the method. They developed a new method of analysis that included learner discussion about the formation of the engagement description. They believed discussion is an important part of the Futurelearn platform design, and it reflects its pedagogical approach, i.e., social constructivism [2].

Definition. *Social constructivism is a learning theory that suggests social interactions* *and cultural activities are the foundation for constructing an individual's knowledge [110]. This theory was proposed by psychologist Lev Vygotsky [111], who developed the concept of the zone of proximal development (ZPD). Rather than providing learners with easy tasks that are within the scope of their knowledge, or complicated tasks that are frustrating, ZPD refers to tasks that a learner can solve with help or guidance from others, e.g., peers or tutors [111]. Vygotsky [111] believed we learn when we interact with our society and environment; thus, we develop. In section 1.3, we summarised Anderson and Dron's view [35] on social constructivism, along with other pedagogical theories regarding MOOC.*

Therefore, Ferguson and Clow's new scoring system focused on three types of content: watching video lectures, completing assessments and commenting in the discussion forums. The scoring system used by Ferguson and Clow [2] was:

- Weekly score = 0: learners did not engage with the course contents.
- Weekly score = 1: learners engaged with a new type of content (e.g., video, text, audio).
- Weekly score = 2: learners posted a comment only, without engaging with new types of content.
- Weekly score = 3: learners posted a comment and engaged with a new type of content.
- Weekly score = 4: learners submitted the assessment late.
- Weekly score = 5: learners engaged with a new type of content and submitted the assessment late.
- Weekly score = 6: learners submitted the assessment late and posted a comment.
- Weekly score = 7: learners engaged with a new type of content, posted a comment and submitted the assessment late.
- Weekly score = 8: learners submitted the assessment on time.
- Weekly score = 9: learners engaged with a new type of content and submitted the assessment on time.

- Weekly score = 10: learners submitted the assessment on time and posted a comment.
- Weekly score = 11: learners engaged with a new type of content, posted a comment and submitted the assessment on time.

Looking at Ferguson and Clow's scoring system [2], we can say they considered all possible types of interaction with the course contents. As for the contents' weight, it is clear that assessment submission had more weight than posting a comment or engaging with a new type of content, and posting a comment had more weight than engaging with a new type of content. As with Kizilcec et al.'s study [1], there is no justification for the contents' weight from Ferguson and Clow; however, we believe it matches the traditional learning in a classroom where passing the exam is more important than participating or attending lectures, and participating is better than only attending lectures.

2.3.2.2 Ferguson and Clow's Calculated Similarities

Ferguson and Clow [2] produced an engagement description for each learner that comprised six or eight variables (for the six-week MOOC and the eight-week MOOC, respectively) and passed these engagement descriptions directly to the K-means clustering as a multidimensional algorithm; they therefore avoided using the L_1 norm equation. They believed that the L_1 norm equation discarded useful information about the learners' weekly interaction [2].

2.3.2.3 Ferguson and Clow's Clustering Algorithm

Ferguson and Clow [2] clustered their data using a K-means clustering algorithm. They used K values of 3 to 8 and measured the Silhouette score for each value; they found that the Silhouette score of K=7 was the maximum [2]. For more information, example 1 on page 34 explains the use of Silhouettes to validate the value of K.

2.3.3 Ferguson and Clow's Findings

Ferguson and Clow's multidimensional clustering [2] resulted in seven clusters, representing seven types of engagements in all four MOOCs; two of these (Sampling and Completing) were found in Kizilcec et al.'s study [1]. Two of the seven engagement types ended up completing the MOOC (see Table 2.4). The seven engagement types were as follows [2]:

1. Samplers: learners joined the course at different stages and engaged with 5% of the course. Few completed one assignment, and 6% to 15% engaged with the discussion. This type was found in all four MOOCs.

2. Strong Starters: learners engaged strongly with the course at the beginning, then dropped out. All completed only one assignment, and 35% to 38% engaged with the discussion. This type was found in all four MOOCs.

3. Returners: learners engaged with the course. Most of them completed the first assignment then returned to complete the second assignment before dropping out. This type was found in three out of four MOOCs, because MOOC 3 had only three assessments.

4. Midway Dropouts: learners engaged with the course and completed three to four assignments before dropping out; 38% to 49% of learners engaged with the discussion. This type was found in MOOCs 1 and 4, whereas MOOCs 2 and 3 did not include this engagement type due to their different course structure.

5. Nearly There: learners engaged with the course and completed most assignments, then dropped out before the end of the course; 48% to 65% of learners engaged with the discussion. This type was found in all four MOOCs.

6. Late Completers: learners engaged with the course and completed the last assignment; however, some assignments were submitted late and some were missed. Forty to 43% of learners engaged with the discussion. The exception was MOOC 3, in which 76% of learners engaged with the discussion. This type was found in all four MOOCs.

7. Keen Completers: learners engaged actively with the course and completed all assignments; 68% to 73% of learners engaged with the discussion. This type was

found in all four MOOCs.

Table 2.4: Engagement types for MOOCs on the Futurelearn platform [2]

Course	Samplers	Strong Starters	Returners	Mid-way Dropouts	Nearly There	Late Completers	Keen Completers
MOOC 1	39%	11%	6%	6%	6%	8%	23%
MOOC 2	39%	14%	8%	-	6%	7%	13%
MOOC 3	37%	8%	-	-	6%	0.2%	7%
MOOC 4	56%	10%	7%	7%	5%	6%	9%

Ferguson and Clow [2] believed that their seven engagement types occurred only after they included learners' interactions with the MOOC discussion forum in their scoring system. Therefore, they argued that pedagogical approaches used to design MOOCs platform should be taken into consideration when analysing learner engagement with MOOCs.

When Ferguson and Clow [2] examined the interaction of learners with the MOOC discussion forum, they found that learners of all engagement types had the same average of posting one comment per learner per week. The exception was the Keen Completers group, where learners posted twice more, with an average of over two comments per learner per week. This allowed Ferguson and Clow [2] to argue that active engagement in a MOOC's discussion was associated with high engagement with the course materials and assessments. Therefore, they recommended additional analysis in the future to explore this relationship further. From their findings, Ferguson and Clow [2] suggested the following points to educators:

- **Previews:** providing a preview of the MOOC's contents and materials might help the Samplers to determine whether to enrol to the MOOC.
- **Weekly Format:** having weekly assessments has given MOOC design a weekly format that might have an impact on a learner's decision to move forward in a course, although it would seem that the decision to move to the next activity would be less significant than moving to the next week. Therefore, changing the course design and linking the contents of the course between the weeks might decrease the dropout rate.

- **Discussion Forum:** Late Completers and Keen Completers made the most use of the discussion forum, which indicates that including a discussion forum in a MOOC's design would help learners to stay engaged and complete the course.

2.3.4 Recommendations

The following points were suggested by Ferguson and Clow [2] after they analysed learner engagement in the Futurelearn platform, to guide future analysis of learner engagement with MOOCs.

1. **Pedagogical Approach:** Ferguson and Clow [2] claimed their work was a strong indication that using a clustering approach that had been used in another MOOC platform would not return a good result. They suggested future research choose an approach that fits the platform context and design.

2. **Hierarchical clustering:** although Ferguson and Clow [2] used K-means as their clustering algorithm, they suggested future research should use a different algorithm for clustering, such as a hierarchical approach, where there is no need for the user to choose a value for K.

3. **Textual clustering:** Ferguson and Clow [2] suggested using a textual clustering approach rather than involving in the process by assigning numerical scores to each engagement description. Assigning numerical scores was needed for K-means, since it uses Euclidean distance to cluster data points. Therefore, Ferguson and Clow suggested using the Levenshtein distance, a distance measurement between texts [112], when using textual clustering.

2.4 Differences Between Kizilcec et al.'s and Ferguson and Clow's Analyses

We conclude this chapter by summarising the key differences between the studies of Kizilcec et al. [1] and Ferguson and Clow [2]. These two fundamental studies used similar methodologies to analyse MOOCs from the Coursera and Futurelearn platforms, respectively.

2.4.1 MOOC Features

The MOOCs examined by Kizilcec et al. [1] and Ferguson and Clow [2] had different features (see Table 2.5). Kizilcec et al.'s data sample comprised three MOOCs from the Coursera platform. The three MOOCs were all computer science courses, albeit not at the same level. The first course was for the high school level, the second was for the undergraduate level, and the third was for the graduate level. On the other hand, Ferguson and Clow's data sample comprised four MOOCs from the Futurelearn platform that represented four different subjects (physical sciences, life sciences, arts and business); they were all designed for beginning learners. Moreover, Ferguson and Clow [2] stated that the majority of their learners were based in the UK, whereas Kizilcec et al. [1] reported having learners from the US, India and Russia. A final difference in the MOOC features was the course structure. The MOOCs in Kizilcec et al.'s work all had the same length and the same assessment periods. However, Ferguson and Clow's MOOCs had different lengths and assessment periods, where one MOOC was two weeks shorter than the other MOOCs, and another MOOC had only three assessments rather than one every week.

Table 2.5: Key differences between the MOOCs from Kizilcec et al.'s and Ferguson and Clow's works

Sample	Kizilcec et al.'s MOOCs	Ferguson and Clow's MOOCs
MOOC subject	Computer Science	Multiple subjects
MOOC level	Three levels	One level
Learners' location	US-based (majority), India, Russia	UK based (majority)
MOOC structure	Same structure	Different structure

2.4.2 Scoring Systems

The scoring systems for the weekly interaction with the MOOCs' contents were different between Kizilcec et al.'s analysis and Ferguson and Clow's analysis. The simple scoring system used by Kizilcec et al. [1], where they aimed to make it adoptable by any MOOC platform, did not show a clear result when used by Ferguson and Clow [2] on a different platform. On the other hand, Ferguson and Clow used a highly detailed scoring system that could only be applied to their platform, Futurelearn; they suggested for future research to do the same and use an approach that is informed by the platform context. Table 2.6 shows the interpretations of the scores used in Kizilcec et al.'s analysis [1], in comparison to Ferguson and Clow's analysis [2].

Table 2.6: Scoring systems used by Kizilcec et al. [1] and Ferguson and Clow [2] to analyse MOOCs from the Coursera and Futurelearn platforms

Score	Interpretation	
	Kizilcec et al. (Coursera platform)	Ferguson and Clow (Futurelearn platform)
0	No participation	No participation
1	Engaged with content, no assessment	Engaged with content, no assessment
2	Late assessment submission	No engagement with content, commented
3	On-time assessment submission	Engaged with content, commented
4		Late assessment submission only
5		Engaged with content, late assessment submission
6		No engagement with content, commented, late assessment submission
7		Engaged with content, commented, late assessment submission
8		On-time assessment submission only
9		Engaged with content, on-time assessment submission
10		No engagement with content, commented, on-time assessment submission
11		Engaged with content, commented, on-time assessment submission

2.4.3 Clustering Method

Another major difference between the studies of Kizilcec et al. [1] and Ferguson and Clow [2] was the way they applied the clustering algorithm. Although both studies used a K-mean clustering algorithm, it was used in different ways. As mentioned in section 2.1.3, the K-means clustering algorithm is a multidimensional algorithm where each data point is represented by two or more points that identify its position in a Euclidean space. The position of each point is important as the algorithm must calculate the Euclidean distance from each point to the cluster's centroid, to determine to which cluster each point belongs.

Kizilcec et al. [1] used K-means as a one-dimensional algorithm, by using the L_1 norm equation to calculate the similarity between learners. The distance matrix was used instead of the Euclidean distance. This step might have caused two learners with different weekly scores to have the same final score and end up in the same cluster. We explain this concern with an example in Section 2.2.2.2.

Ferguson and Clow [2] were concerned about losing useful information when using the L_1 norm equation. Therefore, Ferguson and Clow ran their K-means clustering as a multidimensional algorithm (six- and eight-dimensional) with the full engagement description. This allowed the K-means algorithm to count the weekly engagement of learners in the clustering process.

Chapter 3

Research Methodology used for the Analysis of the Edraak MOOC Platform

Previously, in sections 2.2 and 2.3 we discussed two studies that analysed learners' engagement with two English MOOC from two different platforms, Coursera and Futurelearn [1] [2]. This chapter presents the analytical methods used for our project, which were adapted from those two studies to investigate the natures of learners and their types of engagement on the Arabic MOOC platform Edraak in comparison to the English MOOC platforms Coursera and Futurelearn.

3.1 The Obtained Data From Edraak

To understand the Arabic MOOC learners, we used sets of data generated from the Edraak platform. A MOOC titled Java Programming 1 and presented by the Arab Open University, in Jordan, was provided to us by the Edraak team. This MOOC started on the 4th of October, 2016, and ran for six weeks. The data contained general information about the learners, such as gender, qualification, and nationality. In addition, information about learners' interactions with the MOOC materials was provided, including the number of videos watched, the number of assessments completed, reported grades, and engagement on the discussion forum (i.e. number of comments and replies).

3.1.1 Course Structure

The Java Programming 1 MOOC on Edraak ran for six weeks. Each week, the learners were provided with a number of videos that explained the course contents, and at the end of each week, learners were asked to complete an assessment. No late submission were accepted. Table 3.1 presents the course structure. The first week was an introduction to the course and had only one video to welcome the learners and guide them through downloading a programming platform; there was no assessment. Thus, the second week was the actual commencement of this MOOC.

Table 3.1: Java Programming 1 course structure

Week Number	Number of Videos	Number of Assessments
1	1	0
2	6	1
3	9	1
4	9	1
5	6	1
6	9	1

3.2 Methodology

The first goal of our research was to identify the learners who used the Arabic MOOC platform. The aim was to answer the first research question on page 3, “who are the users of the Arabic MOOC platforms?” Chapter 4 will attempt to answer this question by running some statistical tests to investigate the associations between gender, qualification, human development index (HDI) level, engagement in discussion on the forum, and exposure to the MOOC contents. For these analyses, we performed a chi-squared test when comparing categorical variables, a t-test when comparing two groups of continuous data, and an ANOVA test when comparing more than two groups of continuous data. We performed these tests, using the SPSS software package for statistical analysis, with a confidence interval of 95% ($\alpha = 0.05$).

The second goal of our research was to explore how to answer the question "what are the engagement types of the Arabic-speaking learners?" For this question, we analysed the Edraak learners' data twice. The first analysis adopted the approach from Kizilcec et al. [1] to identify types of learners' engagement on Coursera, with reservations on their 1-dimensional K-means clustering (see section 2.2.2.2). This approach was used only to compare the learners from Edraak and those from Coursera. Our analysis has been presented using the same scoring system and 1-dimensional K-means clustering algorithm used by Kizilcec et al. [1], as mentioned in section 2.2.2.1.

The second analysis of the Edraak data used Ferguson and Clow's approach [2] to identify the types of engagement of Futurelearn learners, and then compare learners from Futurelearn and from Edraak. For this analysis, we did not apply the same scoring system as Ferguson and Clow mentioned in section 2.3.2.1 because our data did not include the learners weekly interaction with the discussion forum. However, we heeded their recommendation to find a unique scoring system to fit the design of the Edraak MOOC. Then we used a multidimensional K-means clustering algorithm to identify Edraak learners' engagement description with MOOCs.

The third goal of our research was to compare the engagement types of Arabic-speaking learners who use Edraak with the English-speaking learners who use Coursera and Futurelearn to answer the question "How is the engagement of Arabic-speaking learners in a MOOC similar to or different from the engagement of English-speaking learners in a MOOC?"

3.2.1 Computing Engagement Descriptions That Represent Learners Weekly Interaction With The MOOC Contents

For the first analysis of the Edraak learners' data, we used the same scoring system as Kizilcec et al. [1], also used to analyse the Coursera data, as mentioned in section 2.2.2.1, which consists of four scores, as follows: (1) When submitting the assessment on time, the score is 3. (2) When submitting the assessment late, the score is 2. (3) When watching videos only and not submitting the assessment, the score is 1. (4) For no engagement at all, the score is 0.

However, since Edraak platform does not allow the late submission of assessments, we eliminated this option, so our system consists of the following three scores:

- Submitting the assessment on time = 3
- Watching videos only, but not submitting the assessment = 2
- No engagement at all = 1

For the second analysis of the Edraak learners' data, we followed Ferguson and Clow's approach [2] used to analyse Futurelearn data, but designed our own scoring system to accommodate the design of Edraak platform. The scoring system used by Ferguson and Clow [2] relied on learners' interactions with MOOC contents that were specific to the Futurelearn platform, such as visiting content, submitting assessments late, submitting assessments on time, and writing comments. As discussed in section 2.3.2.1, Ferguson and Clow covered all possible learner interactions, and they employed a larger scale, with scores ranging from 0 to 11. Thus, our system imitated their approach in that it covered all possible interactions with the MOOC content, but the interactions scored by our system were specific to the Edraak platform.

The structures of Edraak MOOC, mentioned in section 3.1.1, showed only two types of content that could be used to form a scoring system: the number of videos watched per week, and the one assessment at the end of each week. The scoring system allowed us to compute an Engagement Description for each learner, where a score was given to each learner every week to represent their interaction with the course contents. At the end of the course, each learner had an engagement description, which was made of a group of numbers that represented the learners' weekly scores.

From the two types of contents in Edraak, watched videos and assessment, there are four possible scoring systems that can be used. The following example will be used to show the weekly scores and the engagement description (ED) for three learners (learner A, learner B, learner C) in each of the four scoring systems.

Example 2. *A four-week MOOC consists of 4 videos in week one, 5 videos in week two, 4 videos in week three, and 2 videos in week four, with one assessment each week.*
Learner A, watched every video and completed all assessments.
Learner B, watched 4 videos in week one and completed the assessment, watched 2

videos in week two and completed the assessment, and watched 1 video in week three. Learner C, watched 2 videos in week one

First Scoring system: Equal-weighted

Consider the two MOOC contents, videos and assessments, to have equal scoring weight. This is done by assigning one point for each interaction in these MOOC contents. The first scoring system is as follows:

If (video = 0 & assessment = 0) score = 0
If (video = 1 & assessment = 0) score = 1
If (video = 0 & assessment = 1) score = 1
If (video = 1 & assessment = 1) score = 2

The scores range between 0 to 2.

When applying this system to example 2 above:

- Learner A, who watched all videos and completed all assessments receives 5 points in week one, 6 points in week two, 5 points in week three, and 3 points in week four, making the ED = {5,6,5,3}.
- Learner B, who watched 4 videos in week one and did the assessment, watched 2 videos in week two and did the assessment, and watched 1 video in week three receives 5 points in week one, 3 points in week two, 1 point in week three, and no points in week four, making the ED = {5,3,1,0}.
- Learner C, who watched 2 videos in week one, receives 2 points in week one, and no points on weeks two, three and four, making the ED = {2,0,0,0}.

Second Scoring system: Assessment-weighted

The second system is influenced by the scoring system implemented by Kizilcec et al. [1], where assessments have greater weight than other course activities. Here we gave double the score to the assessments. The second scoring system is as follows:

If (video = 0 & assessment = 0) score = 0
If (video = 1 & assessment = 0) score = 1
If (video = 0 & assessment = 1) score = 2
If (video = 1 & assessment = 1) score = 3

The scores range between 0 to 3.

When applying this system to example 2 above:

- Learner A gets 6 points in week one, 7 points in week two, 6 points in week three, and 4 points in week four, making the ED = {6,7,6,4}.
- Learner B gets 6 points in week one, 4 points in week two, 1 point in week three, and no points in week four, making the ED = {6,4,1,0}.
- Learner C gets 2 points in week one and no points in weeks two, three and four, making the ED = {2,0,0,0}.

Third Scoring system: Diligence-weighted

The third system is an attempt at weighing learners' commitment to interacting with all course contents, where the highest score in each MOOC content is awarded to the learners who visit the total number of videos and/or assessments. The third scoring system is as follows, where M is the maximum numbers of videos in each week:

If (video = 0 & assessment = 0) score = 0
If (video < M & assessment = 0) score = 1
If (video = M & assessment = 0) score = 2
If (video = 0 & assessment = 1) score = 3
If (video < M & assessment = 1) score = 4
If (video = M & assessment = 1) score = 5

The scores range between 0 to 5

When applying this system to example 2 above:

- Learner A gets 5 points in week one, 5 points in week two, 5 points in week three, and 5 points in week four, making the ED = {5,5,5,5}.
- Learner B gets 5 points in week one, 4 points in week two, 1 point in week three, and no points in week four, making the ED = {5,4,1,0}.
- Learner C gets 1 point in week one and no points in weeks two, three and four, making the ED = {1,0,0,0}.

Fourth Scoring system: Semi-diligence-weighted

This system goes one step further than the third scoring system (Diligence-weighted), so that learners' scores are weighted according to how much content they visited, meaning that higher numbers of visits results in higher scores. The fourth scoring system is as follows, where M is the maximum numbers of videos in each week:

If (video = 0 & assessment = 0)	score = 0
If (video < 50%M & assessment = 0)	score = 1
If (M > video ≥ 50%M & assessment = 0)	score = 2
If (video = M & assessment = 0)	score = 3
If (video = 0 & assessment = 1)	score = 4
If (video < 50%M & assessment = 1)	score = 5
If (M > video ≥ 50%M & assessment = 1)	score = 6
If (video = M & assessment = 1)	score = 7

The scores range between 0 to 7.

When applying this system to example 2 above:

- Learner A gets 7 points in week one, 7 points in week two, 7 points in week three, and 7 points in week four, making the ED = {7,7,7,7}.
- Learner B gets 7 points in week one, 6 points in week two, 1 point in week three, and no points in week four, making the ED = {7,6,1,0}.

- Learner C gets 2 points in week one and no points in weeks two, three and four, making the ED = {2,0,0,0}.

Table 3.2 shows the interpretations of the four scoring systems. We applied all four scoring systems to analyse the data from Edraak and chose the scoring system based on which system yielded the best clusters from the clustering algorithm.

Table 3.2: The interpretations of the four scoring systems used to analyse Edraak's learners

Score	Scoring Systems			
	Equal-weighted	Assessment-weighted	Diligence-weighted	Semi-diligence-weighted
0	No participation	No participation	No participation	No participation
1	Visit videos OR submit assessment	Visit videos only	Visit some videos only	Visit fewer than half of the videos
2	Visit videos AND submit assessment	Submit assessment only	Visit all videos	Visit half or more of the videos but not all
3		Visit content AND submit assessment	Submit assessment only	Visit all videos
4			Visit some videos, submit assessment	Submit assessment only
5			Visit all videos, submit assessment	Visit fewer than half of the videos, submit assessment
6				Visit half or more of the videos but not all, submit assessment
7				Visit all videos, submit assessment

3.2.2 Calculation of Similarities

For the first analysis of Edraak's learners, we followed Kizilcec et al.'s approach [1], as explained in section 2.2.2.2, and use the L_1 norm equation to calculate the similarities between the learners. In that section, we explained the drawbacks of using the L_1 norm equation; however, we used it in the analysis of our data to imitate the approach used by Kizilcec et al. [1] and to compare our resultant engagement types with theirs.

For the second analysis of Edraak's learners, we will follow Ferguson and Clow's approach [2] and skip the step of calculating similarities between learners, as explained in section 2.3.2. We will input our learners' EDs directly into the K-means clustering as a multidimensional clustering algorithm.

3.2.3 Clustering Algorithm

When clustering Edraak data, using Kizilcec et al.'s approach [1], we will run K-means as a 1-dimensional clustering algorithm. However, when clustering Edraak data following Ferguson and Clow's approach [2], we will run K-means as a 5-dimensional clustering algorithm, where 5 is the number of dimensions in our learners' EDs. The application of K-means in both analyses will be performed using the R programming language and software environment, the full code of which is available in Appendices B and C. To validate the resulting clusters, two steps will be performed.

1. **Select the highest likelihood outcome:** Following Kizilcec et al. [1], performing the clustering operations one hundred times and selecting the outcome having the highest likelihood [1]. Since the first random allocation of centroids may have an effect on the final resulting clusters. For more on the reasoning behind this validation, see the example in section 2.2.2.3.

2. **Determine the optimum number of clusters (K):** Using the NbClust R package that provides 30 indices, where each index suggests a number of clusters for the data set. The NbClust R package selection of the best number of clusters is based on the majority output of the 30 indices. [113]

3. **Measure the cluster cohesion and separation:** Cluster cohesion measure how object in a cluster is closely related, and is measured by the within cluster sum of squares (WSS). Cluster separation measure how a cluster is distinct from other clusters, and is measured by the between cluster sum of squares (BSS). The sum of WSS and BSS is called the total sum of squares (TSS). Measuring the sum

of squares determined the goodness of the clusters resulting from the K-means, where the BSS/TSS ratio close to 1 indicates a good fit. [114]

3.3 Quantitative Analysis

Statistical analysis was performed to answer the research questions, mentioned on page 3:

- Who are the users of the Arabic MOOC platform?
- What are the engagement types of the Arabic-speaking learners?
- How is the engagement of Arabic-speaking learners in a MOOC similar to or different from the engagement of English-speaking learners in a MOOC?

Tables 3.3, 3.4 and 3.5, respectively, show the breakdown of the statistical approach when answering these research questions.

Table 3.3: Statistical approach to answer the first research question

Question	Statistical test
1) Is there a significant difference in qualifications between male and female learners?	Chi-square test significance level $p < 0.05$
2) Is there a significant difference in HDI levels between male and female learners?	Chi-square test significance level $p < 0.05$
3) Is there a significant difference between learner qualifications and HDI levels?	Chi-square test significance level $p < 0.05$
4) Is there a significant difference in the use of discussion forums between male and female learners?	T-test significance level $p < 0.05$
5) Is there a significant difference in the use of discussion forums between differently qualified learners?	ANOVA test significance level $p < 0.05$
6) Is there a significant difference in the use of discussion forums between learners with different HDI levels?	ANOVA test significance level $p < 0.05$

Table 3.4: Statistical approach to answer the second research question

Question	Statistical test
1) Is there a significant difference in the gender proportion among the engagement types?	Chi-square test significance level $p < 0.05$
2) Is there a significant difference in the learners' qualifications among the engagement types?	Chi-square test significance level $p < 0.05$
3) Is there a significant difference in the distribution over countries with different HDI levels among the engagement types?	Chi-square test significance level $p < 0.05$
4) Are there significant differences in the number of comments among the engagement types?	ANOVA test significance level $p < 0.05$
5) Are there significant differences in the exposure to the course materials among the engagement types?	ANOVA test significance level $p < 0.05$

Table 3.5: Statistical approach to answer the third research question

Question	Statistical test
1) Is there a significant difference in the engagement types proportion between Arabic- and English-learners?	Chi-square test significance level $p < 0.05$
2) Is there a significant difference in the gender proportion between Arabic- and English-learners?	Chi-square test significance level $p < 0.05$
2.1) Is there a significant difference in the gender proportion between Arabic- and English-learners engagement types?	Chi-square test significance level $p < 0.05$
3) Is there a significant difference in the distribution over countries with different HDI levels between Arabic- and English-learners?	Chi-square test significance level $p < 0.05$
3.1) Is there a significant difference in the distribution over countries with different HDI levels between Arabic- and English-learners engagement types?	Chi-square test significance level $p < 0.05$
4) Is there a significant difference in the use of discussion forums between Arabic- and English-learners?	Chi-square test significance level $p < 0.05$

3.3.1 Chi-square Test

Chi-Square is a statistical test used for analysing categorical data by comparing the expected count with the actual count of the sample in each category. The expected count is the expected frequencies in each cell if the null hypothesis is true, and can be calculated using equation 3.1, where e is the expected count, $\sum r$ is the sum of data in the row, $\sum c$ is the sum of data in the column, and n is the total number of data [115].

$$e = \frac{\sum r \times \sum c}{n} \tag{3.1}$$

Chi-Square test compares the expected count with the actual count with a significance level of $p < 0.05$ using equation 3.2, where X^2 is the Chi-Square value, a is the actual count of data in a category, and e is the expected count of data in a category [115].

$$X^2 = \sum \frac{(a-e)^2}{e} \tag{3.2}$$

3.3.1.1 Adjusted residual

Similar to ANOVA test, the significant in Chi-Square indicates that a different existed between the expected counts and the actual counts of the sample categories, however, it does not clearly indicates which one is different. Therefore the adjusted residual value of ± 1.96 is used to determine the significant difference between categories. When the adjusted residual is ≥ 1.96 or ≤ -1.96 in a category, then the obtained difference is significant. The adjusted residual can be calculated using equation 3.3, where a is the actual count of data in a category, and e is the expected count of data in a category, r is the row value, c is the column value, and n is the total number of data [115].

$$\text{Adjusted residual} = \frac{(a-e)}{\sqrt{e(1-\frac{r}{n})(1-\frac{c}{n})}} \tag{3.3}$$

Moreover, if an adjusted residual value is ≥ 1.96 then the category is over-represented, meaning that the actual count is significantly more that the expected count. On the other hand, if an adjusted residual value is ≤ 1.96 then the category is under-represented, meaning that the actual count is significantly less that the expected count [115].

3.3.1.2 Strength of Chi-Square Significance

To measure the strength of the significance obtained from Chi-Square test, Cramer's V test produces a value between 0 and 1, which is used to determine whether the strength of Chi-Square significant is "Large","Medium" or "Low" [116]. Low significant is usually resulted due to large sample size, where any small change in the sample is detected as significant result by Chi-Square test. This means that significant differences with “Large” and “Medium” strength are considered acceptable, whereas significant differences with “Low” strength are considered not acceptable [116].

To determine the strength of Chi-Square significance, Cramer’s V value is used with degree of freedom (df) value. The df value is produced by Chi-Square test, which gives an indication of the number of compared categories. The biggest the df value, the larger number of compared categories. df is calculated by multiplying $(R-1)$ and $(C-1)$, where R and C are the number of rows and columns, respectively [115].

The obtained df and Cramer’s V values are interpreted from table 3.6 to determine the level of strength of Chi-Square significance [116]. This table can be used for df values that range between 1 to 5. For df values larger than 5, only Cramer’s V value is used to determine the strength of significance by converting it to “Cohen’s w” value using equation 3.4, where V is Cramer’s V value and R is number of rows. The strength of significance levels of “Small”, “Medium” and “Large” have a Cohen’s w values of 0.10, 0.30, and 0.50 respectively.

$$w = V\sqrt{R-1} \tag{3.4}$$

Table 3.6: Interpreting the strength of significance levels from df values

df	Strength of significance levels		
	Small	Medium	Large
1	0.1	0.3	0.5
2	0.07	0.21	0.35
3	0.06	0.17	0.29
4	0.05	0.15	0.25
5	0.05	0.13	0.22

3.3.2 T-test

T-test is a statistical test used for comparing the Means of two groups of continues data, that is normally distributed. A difference between two groups is consider significant when T-test report a p value less than 0.05. Prior to using T-test, the following assumptions should exist in the data sample [115][117]:

- The data should follows a continuous or ordinal scale.
- The data should be simple and randomly selected from the total population.
- The data should be normally distribution.
- The data size should be reasonably large.

Equation 3.5 shows the formula of T-test, where X_1 is Mean of the first group, X_2 is Mean of the second group, n_1 is the sample size of the first group, n_2 is the sample size of the second group, and S_p is the pooled standard deviation which is calculated using equation 3.6 where s_1 is the standard deviation of the first group and s_2 is the standard deviation of the second group.

$$t = \frac{X_1 - X_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{3.5}$$

$$S_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \tag{3.6}$$

3.3.3 One-way ANOVA Test

ANOVA is a statistical test used for comparing the Means of more than two groups of continues data. This is by generalising the t-test to more than two groups. ANOVA requires the same assumptions of t-test [115].

A significant deference between the compared groups is determined by a p value less than 0.05. This significant indicates that a different existed between the compared groups, however, ANOVA does not clearly indicates where the significant difference is from the compared groups. Therefore, post tests are required to identify where the significant occur [115].

Determining which post test to use depends on the variances within the compared groups. If the groups variances are NOT equal then the suggested post test is “Games-Howell”. However, if the variances are equal then determining the post test depends on the the sample size. Comparing groups with equal sample size suggests using “Tukey’s HSD” post test. On the other hand, Comparing groups with diffrent sample size suggests using “Gabriel” or “Hochberg’s GT2” post tests.

Chapter 4

MOOC Learners in Edraak

In Chapter 3, we explained the methodology of the analyses we applied to the research questions. This chapter focuses on the first research question (Who are the users of the Arabic MOOC platform?) to identify the general features of Edraak learners. This question was answered by performing statistical analysis and discussing some general observations of Edraak's data.

Edraak Data

Our analysis of Edraak learners used a computer science MOOC titled "JAVA Programming 1" that targeted beginner level learners. The data in this MOOC consists of 2,736 Arabic-speaking learners. Table 4.1 summarises the demographics of these learners, showing the number of male learners to be more than double the number of females (68% to 32%). Moreover, The majority of learners hold a Bachelor's or a high school degree, with percentages of 43.3% and 32.2%, respectively.

Table 4.1 also shows the distribution of learners over the 2018 Human Development Index (HDI), which classify the development of countries based on life expectancy, education and income [118]. More on the HDI and its calculations can be found in Appendix A. Most learners comes from countries that have a medium level of HDI (50.7%), while, the low level of HDI represents the smallest proportion (9.4%).

Table 4.1: Edraak's MOOC demographics

	Java Programming 1 course	
Gender		
Males	1861	(68%)
Females	872	(31.9%)
Not Determined	3	(0.1%)
Qualifications		
No formal education	19	(0.69%)
Junior high school	136	(4.97%)
High school	881	(32.2%)
Associate	97	(3.54%)
BSc	1184	(43.27%)
MSc	313	(11.44%)
PhD	89	(3.25%)
Other	257	(9.4%)
HDI level		
Very High	396	(14.5%)
High	662	(24.2%)
Medium	1387	(50.7%)
Low	257	(9.4%)
Total number of comments	39	
Total learners enrolled	2736	

To analyse the data in table 4.1 in a way that answers the first research question using the provided Edraak MOOC demographics (gender, qualification, HDI level and number of comments), we applied the following statistical approach:

1. Is there a significant difference in qualifications between male and female learners? If so, what is the strength of this significant?
2. Is there a significant difference in HDI levels between male and female learners? If so, what is the strength of this significant?
3. Is there a significant difference between learner qualifications and HDI levels? If so, what is the strength of this significant?

4. Is there a significant difference in the use of discussion forums between male and female learners?

5. Is there a significant difference in the use of discussion forums between differently qualified learners?

6. Is there a significant difference in the use of discussion forums between learners with different HDI levels?

Questions 1 to 3 of the approach will measure significance using the **Chi-square** test, while question 4 will measure it with a **T-text**, and questions 5 and 6 will use an **ANOVA** statistical test. For more information about Chi-Square, t-test and ANOVA tests, see sections 3.3.1, 3.3.2 and 3.3.3 respectively. The significance levels considered in these tests are in table 4.2.

Table 4.2: The considered significance levels for the used statistical tests

Statistical Tests	Significance Level
Chi-Square	$p < 0.05$
Adjusted Residuals	± 1.96
Cramer's V	See table 3.6 in section 3.3.1.2
T-test	$p < 0.05$
ANOVA	$p < 0.05$

4.1 Qualification Differences Between Males and Females in Edraak MOOC

Figure 4-1 shows the proportions of Edraak learner qualifications for males and females that range from "No formal education" to "PhD degree". In general, males and females have similar proportions in all educational levels. The Chi-Square test (Table 4.3) shows a significant difference ($X^2 = 20.644$, $df = 6$, $p < 0.01$) between males and females in the proportions of the (BSc) and (Other) categories. However, Cramer's V test shows that the strength of this significance is **Small** with an unacceptable value of 0.09 [116]. Therefore, these data suggest that the difference in the qualifications between male and female Edraak learners is **insignificant**.

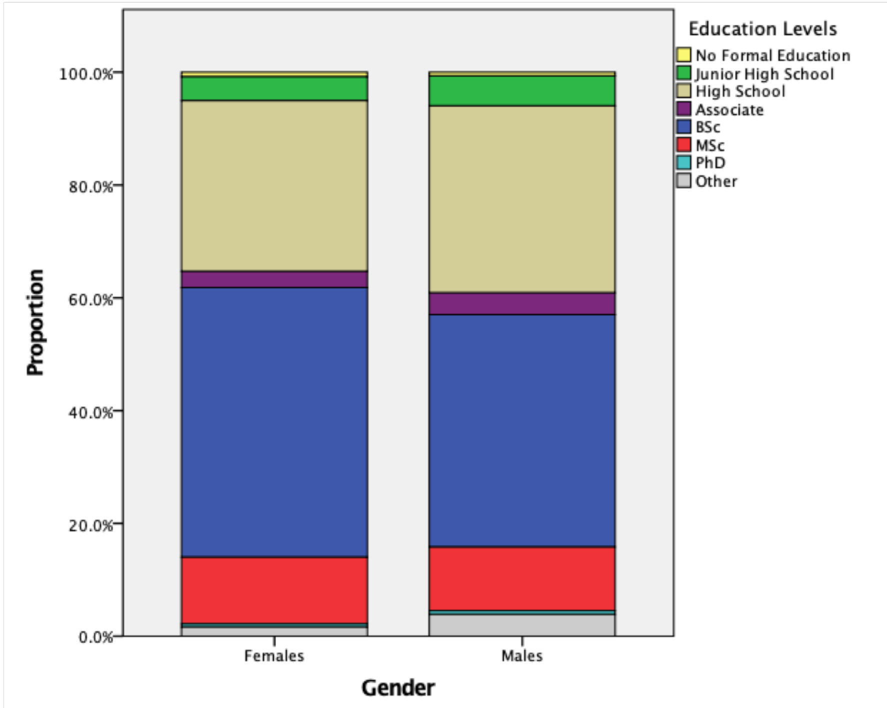


Figure 4-1: The proportions of males and females qualifications in Edraak MOOC

Table 4.3: Crosstabulation of qualifications and gender of Edraak learners

		Females	Males	Total
No Formal Education	Count	7	12	19
	Expected Count	6.1	12.9	
	Adjusted Residual	0.5	-0.5	
Junior High School	Count	37	99	136
	Expected Count	43.4	92.6	
	Adjusted Residual	-1.2	1.2	
High School	Count	264	617	881
	Expected Count	281.1	599.9	
	Adjusted Residual	-1.5	1.5	
Associate	Count	25	72	97
	Expected Count	30.9	66.1	
	Adjusted Residual	-1.3	1.3	
BSc	Count	417	767	1184
	Expected Count	377.8	806.2	
	Adjusted Residual	3.2	-3.2	
MSc	Count	103	210	313
	Expected Count	99.9	213.1	
	Adjusted Residual	0.4	-0.4	
PhD	Count	5	12	17
	Expected Count	5.4	11.6	
	Adjusted Residual	-0.2	0.2	
Other	Count	14	72	86
	Expected Count	27.4	58.6	
	Adjusted Residual	-3.2	3.2	
Total		872	1861	2733

Remark. *Among all the qualifications of Edraak MOOC learners, Bachelor's degree holders represent the majority. Comparing the proportions of male and female learners in each qualification shows no significant difference.*

4.2 The Proportions of Male and Female Edraak MOOC Learners in Terms of HDI Levels

We explored the HDI level proportions of the Edraak MOOC learners to determine if differences existed between male and female learners at each level. Such study grants a better understanding of how a country's development status affects learner engagement in MOOCs.

Figure 4-2 shows the proportions of male and female learners in each HDI level of Edraak learner countries. Generally, the highest proportion of male and female learners are from countries with a Medium HDI level, whereas the smallest proportion of learners came from countries with Low HDI level. A Chi-Square test (Table 4.4) shows a **significant** difference between males and females in their HDI levels ($X^2 = 115.902$, $df = 3$, $p < 0.001$). The adjusted residual values (Table 4.4) showed a significant difference between males and females in the very high, medium, and low HDI levels. The female proportion is over-represented at the very high HDI level, while the male proportion is over-represented at both medium and low HDI levels. Cramer's V test indicates that the strength of these significant differences is **Medium** with an acceptable value of 0.2 [116].

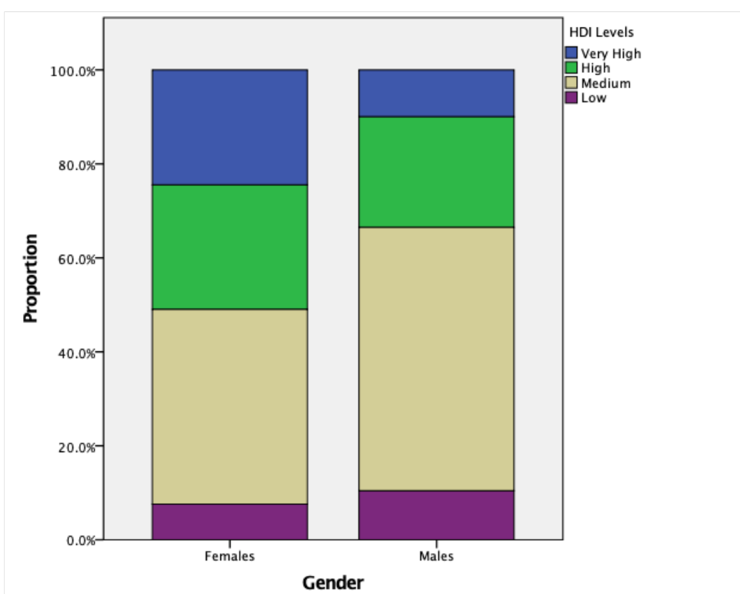


Figure 4-2: The proportions of Edraak males and females in the HDI levels of their countries

Table 4.4: Crosstabulation of countries HDI levels and gender of Edraak learners

		Females	Males	Total
Very High	Count	212	183	395
	Expected Count	126.8	268.2	
	Adjusted Residual	9.9	-9.9	
High	Count	230	432	662
	Expected Count	212.5	449.5	
	Adjusted Residual	1.7	-1.7	
Medium	Count	359	1028	1387
	Expected Count	445.2	941.8	
	Adjusted Residual	-7.1	7.1	
Low	Count	66	191	257
	Expected Count	82.5	174.5	
	Adjusted Residual	-2.3	2.3	
Total		867	1834	2701

Remark. *The majority of male and female Edraak learners, are from countries with a Medium HDI level. Comparing the proportions of male and female learners in all HDI levels showed significant differences, with the Very High HDI level countries having a higher female proportion and the Medium and Low HDI countries having higher male proportions.*

4.3 Edraak Learners Qualifications and The HDI Levels of Their Countries

Overall, learners with Bachelor and high school degrees represent the highest proportions of enrolment in Edraak MOOC qualifications, as seen in Figure 4-3. Notably, High HDI level countries have fewer learners with high school qualifications and more learners with Master's degrees compared to other HDI level countries. In addition, Low HDI level countries have the highest proportion of learners with Associate's degrees. That being said, the Chi-Square test displayed in Table 4.5 shows a significant difference between the compared categories ($X^2 = 158.012$, $df = 21$, $p < 0.001$), but the strength of this significance has an unacceptable **Small** value of 0.14 [116]. This result suggests that the differences between the qualifications of Edraak MOOC learners at various HDI levels is **insignificant**.

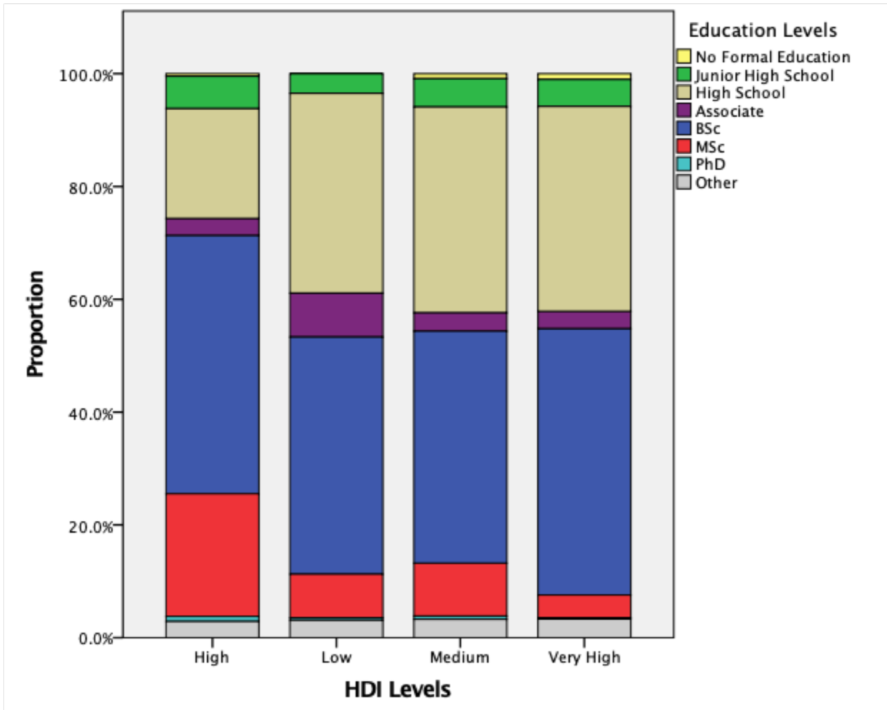


Figure 4-3: The proportions of Edraak learners qualifications in the HDI levels of their countries

Table 4.5: Crosstabulation of qualifications and countries HDI levels of Edraak learners

		Very High	High	Medium	Low	Total
No Formal Education	Count	4	3	12	0	19
	Expected Count	2.8	4.7	9.8	1.8	
	Adjusted Residual	0.8	-0.9	1	-1.4	
Junior High School	Count	19	38	70	9	136
	Expected Count	19.9	33.3	69.8	12.9	
	Adjusted Residual	-0.2	1	0.03	-1.2	
High School	Count	144	129	506	91	870
	Expected Count	127.5	213.2	446.6	82.7	
	Adjusted Residual	1.9	-8.1	4.9	1.2	
Associate	Count	12	20	45	20	97
	Expected Count	14.2	23.8	49.8	9.2	
	Adjusted Residual	-0.6	-0.9	-1	3.8	
BSc	Count	187	303	571	108	1169
	Expected Count	171.3	286.4	600.1	111.2	
	Adjusted Residual	1.7	1.5	-2.3	-0.4	
MSc	Count	16	144	130	20	310
	Expected Count	45.4	76	159.1	29.5	
	Adjusted Residual	-5	9.6	-3.5	-2	
PhD	Count	1	6	8	1	16
	Expected Count	2.3	3.9	8.2	1.5	
	Adjusted Residual	-1	1.2	-0.1	-0.4	
Other	Count	13	19	45	8	85
	Expected Count	12.5	20.8	43.6	8.1	
	Adjusted Residual	0.2	-0.5	0.3	0.03	
Total		396	662	1387	257	2702

Remark. *Most of Edraak learners are Bachelor and high school degree holders. Comparing the proportions of Edraak learner qualifications in the HDI levels of their countries shows no significant difference.*

4.4 Discussion Forum Participation Among The Demographic Categories of Edraak Learners

The Edraak platform provides a discussion forum that allows enrolled learners to discuss the MOOC contents and encourages collaborative learning [119]. In general, the participation in the discussion forum is minimal when compared to the number of enrolled learners (39 comments to 2,736 learners, as shown in Table 4.1). We wanted to know whether there is a significant difference between the demographic categories (gender, qualifications and countries HDI levels) of Edraak learners in the discussion forum participation. This could contribute in defining the features of Edraak learners.

However, the low number of comments violations the assumption of most statistical tests ,such as T-test and ANOVA, of using a reasonably large sample size [115]. Therefore, we can not studying the distribution profile of comments within the different demographic categories (gender, qualifications and country HDI levels) of Edraak learners. We can only mention that, despite the small participation of Edraak learners, males have more comments on the discussion forum than females (35 to 4).

4.5 Discussion

In this chapter, we used data from a computer science MOOC on the Edraak platform (titled: JAVA Programming 1), which was designed for beginner level learners. This MOOC platform delivers its courses in the Arabic language and targets Arabic learners. The data provided by the Edraak team contains learner demographic information that include: gender, qualification and countries. The data also contains the number of comments made by each learner on the discussion forum during the course of the MOOC, and information about the learners' weekly interaction with the MOOC contents that include: video lectures and assessments. We wanted to understand **Who are the users of the Arabic MOOC platform?** to see if we could identify the major features of Edraak learners. Answers were obtained using Edraak learner demographic information.

Our results showed that Edraak MOOC attracts Arabic learners with Bachelor's and high school degrees more than other qualifications. This might be influenced by the level of the chosen MOOC, as it was designed for beginner learners, but this should be investigated further at a qualitative level. The results also showed that learners from Medium HDI level countries represented the highest proportion of learners in the Edraak MOOC, whereas the smallest fraction of learners came from countries with a Low HDI level. Comparing the gender categories, the majority of learners enrolled from Very High HDI level countries were females, whereas males represented a majority in both Medium and Low HDI level countries. This finding raises the question of what influences the interest of MOOCs in Medium HDI level countries. MOOC accessibility may have affected the enrolment of learners in Low HDI countries. Finally, we observed that males participated more than females in the discussion forum, however the total participation of Edraak MOOC learners was very low. This result might interest MOOC designers and providers who believe that learners communication in MOOCs is important for their collaborative learning [120]. Table 4.6 summarise the results of the statistical analyses.

By studying the users of this Arabic MOOC platform, we observed that males represent the majority of Edraak learners within gender category. Learners with Bachelor's degree have the highest proportion within the qualification category followed by high school degree. The majority of Edraak learners come from countries with Medium HDI level. Finally, Edraak MOOC learners have a poor use of the discussion forum. These observations represent the main features of Edraak MOOC learners, which might benefit Edraak designers to better understand their learners and improve their MOOCs to suit learner needs.

Table 4.6: Summary of the statistical analyses

Questions	Statistical test	Significance $p < 0.05$	Strength of significance (Cramer's V)
1. Is there a significant difference in qualifications between male and female learners?	Chi-Square	✔	Small
2. Is there a significant difference in HDI levels between male and female learners?	Chi-Square	✔	Medium
3. Is there a significant difference between learner qualifications and HDI levels?	Chi-Square	✔	Small

Chapter 5

One-dimensional K-means Clustering on The Edraak Data

As mentioned in the methodology in Chapter 3, we will analyse the Edraak data using the two approaches that were used to analyse the Coursera and Futurelearn MOOCs [1] [2]. In this chapter we start with Coursera's approach that used the 1-dimensional K-means clustering algorithm [1], see Section 2.2.2.1 for details of this approach. The purpose of this analysis is to answer the second research question: **What are the engagement types of the Arabic-speaking learners?**

In Section 5.1 we follow their "Engagement Description" computation process, and we use the L_1 norm to calculate similarities between learners. Section 5.2 presents Edraak's Engagement Types using the 1-dimensional K-means clustering algorithm. Section 5.3 concludes this chapter with the quantitative statistical analysis that compared the resulting engagement types.

5.1 Edraak's Engagement Description: Kizilcec et al. Method

To form an Engagement Description, Kizilcec et al. [1] assigned a weekly score to each learner that represents their interaction with the MOOC contents. Their scoring system was mainly depended only on completing the weekly assessment regardless of the completion of the other MOOC content. In case a learner did not interact with the weekly assessment, watching video lectures would be used instead. Kizilcec et al. [1] used a weekly scoring system that assigned learners **3** points if they submitted the assessment on time, **2** points if they submitted the assessment late, **1** point if they missed the assessment but watched video lectures, and a score of **0** if they did not participate at all.

Following the Kizilcec et al. [1] scoring system, every week we assigned scores to Edraak's learners based on their interactions with the MOOC content. However, since the Edraak platform doesn't allow the late submission of assessments, we did not include this case, making our scoring system as follows:

- Submitting assessment on time = 3
- Watching videos only, but not submitting the assessment = 2
- No engagement at all = 1

At the end of the MOOC, each learner had an engagement description consisting of five numbers, which represented the scores they got every week. Then, we applied the L_1 norm equation to the learners' engagement descriptions and used the result of this calculation as an input for the 1-dimensional K-means clustering algorithm (more about the use of L_1 norm equation is detailed is Section 2.2.2.2). We explored K-means clustering with K value from 1 to 10. The majority of indices in the NbClust R package suggested that K=3 is the best number of clusters. In addition, the BSS/TSS ration of 0.9 showed a good fit at K=3, see Figure 5-1. For more information about the validation tests, see Section 3.2.3. The K-means clustering, and the validation tests were performed using the R programming language and software environment. The code can be found in Appendix B.

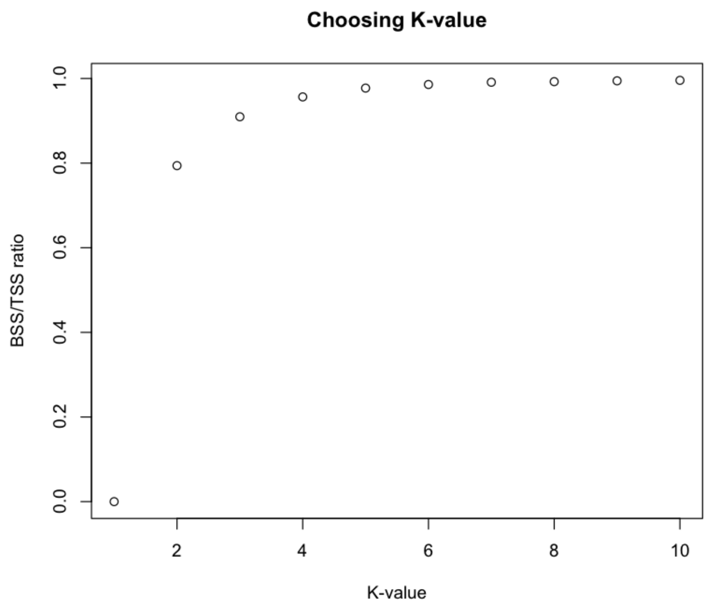


Figure 5-1: Choosing the optimum K value for Edraak data K-means clustering from the BSS/TSS ratio. The data was clustered multiple times with K = 1:10

> **Remark.** *The scoring system used to compute the Edraak's engagement description was: Submitting an assessment on time resulted in a score of 3; watching video lectures but not submitting the assessment resulted in a score of 2; and no engagement at all resulted in a score of 1.*
> *Clustering Edraak data was performed using a 1-dimensional K-means clustering algorithm, with K= 3.*

5.2 Edraak's Engagement Types: Kizilcec et al. Method

Using a 1-dimensional K-means clustering algorithm in classifying Edraak learners resulted in three clusters (engagement types). To define these three clusters of learners, we analysed their interaction patterns with the contents throughout the MOOC.

Cluster 1: This was the largest cluster with 2302 learners out of 2736. At the beginning of the MOOC, more that half of the learners did not watch any videos, and 95% did not submit the weekly assessment. This sharply declined in the following week, where the vast majority of learners did not watch any videos, and none of them submitted the weekly assessment. By the end of the course, learners continued to not interact with the course materials, see Figure 5-2. In general, learners in this cluster stopped engaging with the MOOC after the first week. Learners engagement pattern in this cluster was similar to the "Sampling" cluster of Coursera [1], therefore, we called this engagement type **Sampling**. For more on the detailed engagement of Cluster 1, see Tables 2 and 3 in Appendix D.

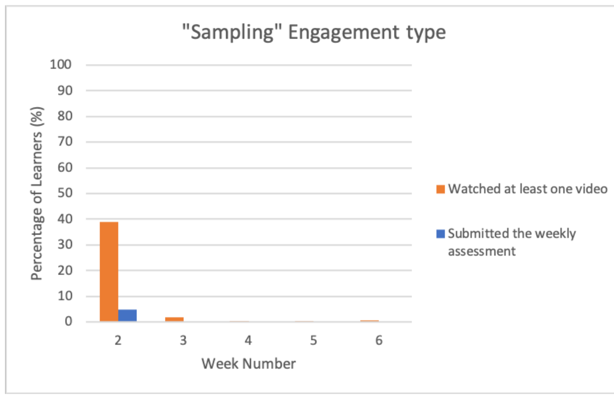


Figure 5-2: Learners interaction with Edraak MOOC contents in the Sampling engagement type

Cluster 2: There were 254 learners out of 2736 in the second cluster. At the beginning of the MOOC, the majority of learners watched multiple videos and submitted the assessment. The high learner engagement with videos continued in the following week, however, the percentage of learner assessment submissions was dropped by more than half. Learners interaction with the MOOC contents continued to decrease in the following weeks, see Figure 5-3. In general, learners in this cluster started the MOOC with good engagement, and then appeared to lose interest midway through the course. This engagement profile is similar to the "Disengaging" cluster of Coursera [1], therefore, we called this engagement type **Disengaging**. For more on the detailed engagement of Cluster 2, see Tables 4 and 5 in Appendix D.

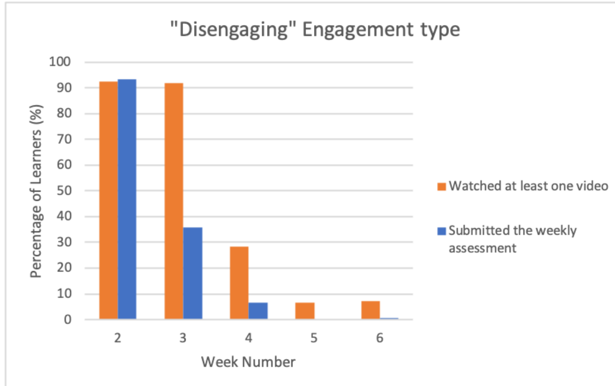


Figure 5-3: Learners interaction with Edraak MOOC contents in the Disengaging engagement type

Cluster 3: The third cluster had 180 learners out of 2736. At the beginning, over 60% of the learners engaged with the videos, and almost all of them submitted the assessment. Midway through the course, more than half of the learners stopped engaging with the videos, but the vast majority of them continued to submit the assessment. By the end of the course, most learners continued not watching the videos, However around 80% of them submitted the assessment, see Figure 5-4. In general, learners in this cluster were interested in completing the course, even though they were not consistent in engaging with the videos. Therefore, this cluster was called **Completing**. For more on the detailed engagement of Cluster 3, see Tables 6 and 7 in Appendix D.

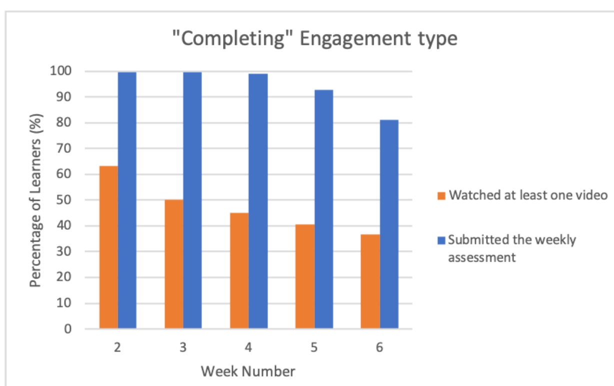


Figure 5-4: Learners interaction with Edraak MOOC contents in the Completing engagement type

Remark. *The K-means clustering algorithm resulted in three engagement types: "Sampling", which consisted of learners who dropped out early in the MOOC; "Disengaging", which consisted of learners who dropped out in the middle of the MOOC; "Completing", which consisted of learners who completed the MOOC.*

5.3 Quantitative Analysis

The K-means clustering algorithm, in Section 5.2, resulted in three engagement types of Edraak learners. There were **Sampling** learners, who had a low interaction with the MOOC contents, video lectures and assessment, and dropped out early in the course, see Figure 5-2. **Disengaging** learners initially interacted with the MOOC contents, but dropped out in the middle of the course, see Figure 5-3. Finally, there were **Completing** learners, who consistently interacted with the MOOC contents, and completed the course, see Figure 5-2. To further understand Edraak learners, this section compared the demographics of these engagement types in Tables 5.1, 5.2 and 5.3. This might help in finding potential reasons behind the identified engagement types.

Table 5.1: Gender demographics of Edraak's learners within the engagement types (Using 1-dimensional K-means clustering)

Gender	Engagement Types			Total
	Sampling	Disengaging	Completing	
Males	1546	177	138	1861
% within gender	83.08%	9.51%	7.41%	100%
% within engagement type	67.15%	69.68%	76.66%	-
Females	753	77	42	872
% within gender	86.35%	8.83%	4.82%	100%
% within engagement type	32.71%	30.31%	23.33%	-
Not specified	3	-	-	3
% within gender	100%	-	-	100%
% within engagement type	0.13%	-	-	-
Total	2302	254	180	2736
% within gender	84.14%	9.29%	6.57%	100%
% within engagement type	100%	100%	100%	-

Table 5.2: Qualification demographics of Edraak's learners within the engagement types (Using 1-dimensional K-means clustering)

Qualification	Engagement Types			Total
	Sampling	Disengaging	Completing	
No formal education	19	0	0	19
% within qualification	100%	0%	0%	100%
% within engagement type	0.83%	0%	0%	-
Junior high school	113	17	6	136
% within qualification	83.09%	12.5%	4.41%	100%
% within engagement type	4.91%	6.7%	3.34%	-
High school	728	91	62	881
% within qualification	82.63%	10.33%	7.04%	100%
% within engagement type	31.63%	35.83%	34.44%	-
Associate	81	11	5	97
% within qualification	83.51%	11.34%	5.15%	100%
% within engagement type	3.52%	4.33%	2.78%	-
BSc	999	106	79	1184
% within qualification	84.37%	8.96%	6.67%	100%
% within engagement type	43.4%	41.73%	43.89%	-
MSc	277	20	16	313
% within qualification	88.5%	6.39%	5.11%	100%
% within engagement type	12.03%	7.87%	8.89%	-
PhD	13	0	4	17
% within qualification	76.47%	0%	23.53%	100%
% within engagement type	0.56%	0%	2.22%	-
Other	72	9	8	89
% within qualification	80.9%	10.11%	8.99%	100%
% within engagement type	3.13%	3.54%	4.44%	-
Total	2302	254	180	2736
% within qualification	84.13%	9.29%	6.58%	100%
% within engagement type	100%	100%	100%	-

Table 5.3: HDI level of Edraak's learners countries within the engagement types (Using 1-dimensional K-means clustering)

HDI level	Engagement Types			Total
	Sampling	Disengaging	Completing	
Very High	327	36	33	396
% within HDI level	82.58%	9.09%	8.33%	100%
% within engagement type	14.20%	14.17%	18.33%	-
High	543	74	45	662
% within HDI level	82.02%	11.18%	6.8%	100%
% within engagement type	23.58%	29.13%	25%	-
Medium	1185	124	78	1387
% within HDI level	85.43%	8.94%	5.63%	100%
% within engagement type	51.47%	48.81%	43.33%	-
Low	218	17	22	257
% within HDI level	84.82%	6.62%	8.56%	100%
% within engagement type	9.47%	6.69%	12.22%	-
Total	2302	254	180	2736
% within HDI level	84.14%	9.28%	6.58%	100%
% within engagement type	100%	100%	100%	-

To compare the demographics of the engagement types of Edraak learners, the following statistical approach was used:

1. Is there a significant difference in the gender proportion among the engagement types? If so, what is the strength of this significant?

2. Is there a significant difference in the learners' qualifications among the engagement types? If so, what is the strength of this significant?

3. Is there a significant difference in the distribution over countries with different HDI levels among the engagement types? If so, what is the strength of this significant?

4. Are there significant differences in the number of comments among the engagement types?

5. Are there significant differences in the exposure to the course materials among the engagement types?

The **chi-square** test was used to analyse the first three questions, where the compared data were categorical. Questions 4 and 5 were analysed using the **ANOVA** test, where the compared data were continuous (for more information about the chi-squared and ANOVA tests, see Sections 3.3.1 and 3.3.3 respectively). The significance levels considered in these tests are shown in Table 5.4.

Table 5.4: The considered significance levels for the used statistical tests

Statistical Tests	Significance Level
Chi-Square	$p < 0.05$
Adjusted Residuals	± 1.96
Cramer's V	See table 3.6 in section 3.3.1.2
ANOVA	$p < 0.05$

5.3.1 Gender Proportions in the Engagement Types of Edraak

The data shows that males proportion appears to be higher than females proportion in all three engagement types, see Figure 5-5. The chi-squared test showed a significant difference in the gender proportions between the engagement types of Edraak ($X^2 = 7.144, df = 2, p < 0.05$), see Table 5.5. However, the Cramer's V test reported that the strength of this significance was **Small** with a value of 0.05, which is considered **not acceptable** [116].

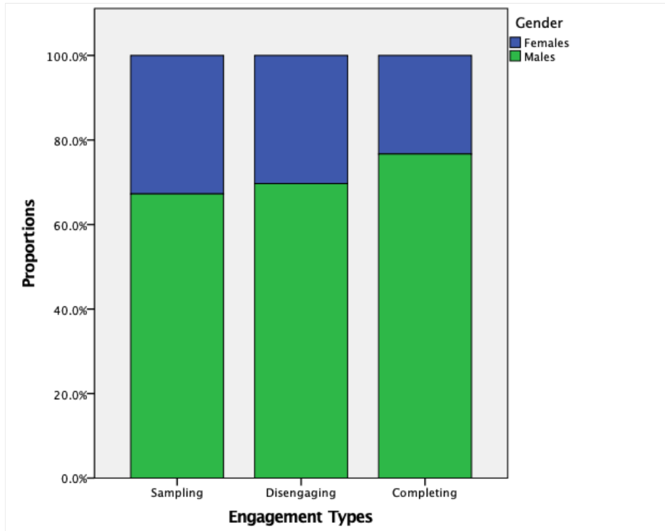


Figure 5-5: The gender proportions within the engagement types of Edraak learners

Table 5.5: Crosstabulation of gender proportions and the engagement types of Edraak learners

		Sampling	Disengaging	Completing	Total
Males	Count	1546	177	138	1861
	Expected Count	1565.5	173	122.6	1861
	% within Engagement Types	67.2%	69.7%	76.7%	68.1%
	Adjusted Residual	-2.2	0.6	2.6	
Females	Count	753	77	42	872
	Expected Count	733.5	81	57.4	872
	% within Engagement Types	32.8%	30.3%	23.3%	31.9%
	Adjusted Residual	2.2	-0.6	-2.6	
Total	Count	2299	254	180	2733
	% within Engagement Types	100%	100%	100%	100%

5.3.2 Qualification Proportions in the Engagement Types of Edraak

The qualification proportions of Edraak learners appears to be equally distributed between the three engagement types, figure 5-6. This observation was supported by the chi-square test (see Table 5.6) that showed **no significance** between the compared groups ($X^2 = 23.190, df = 14, p > 0.05$). Generally, high school and bachelor's degree holders represented the majority of learners in all three engagement types.

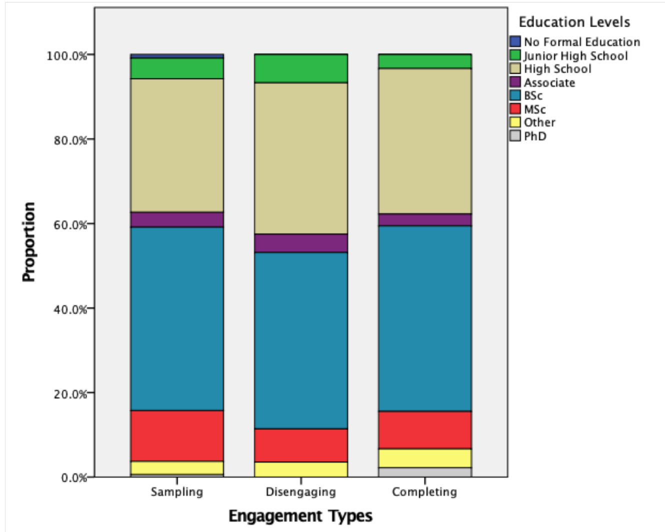


Figure 5-6: The qualification proportions within the engagement types of Edraak learners

Table 5.6: Crosstabulation of qualification proportions and the engagement types of Edraak learners

		Sampling	Disengaging	Completing	Total
No formal education	Count	19	0	0	19
	Expected Count	16	1.8	1.3	19
	% within Engagement Types	0.8%	0%	0%	0.7%
	Adjusted Residual	1.9	-1.4	-1.2	
Junior high school	Count	113	17	6	136
	Expected Count	114.4	12.6	8.9	136
	% within Engagement Types	4.9%	6.7%	3.3%	5%
	Adjusted Residual	-.3	1.3	-1	
High school	Count	728	91	62	881
	Expected Count	741.3	81.8	58	881
	% within Engagement Types	31.6%	35.8%	34.4%	32.2%
	Adjusted Residual	-1.5	1.3	0.7	
Associate	Count	81	11	5	97
	Expected Count	81.6	9	6.4	97
	% within Engagement Types	3.5%	4.3%	2.8%	3.5%
	Adjusted Residual	-0.2	0.7	-0.6	
BSc	Count	999	106	79	1184
	Expected Count	996.2	109.9	77.9	1184
	% within Engagement Types	43.4%	41.7%	43.9%	43.3%
	Adjusted Residual	0.3	-0.5	0.2	
MSc	Count	277	20	16	313
	Expected Count	263.4	29.1	20.6	313
	% within Engagement Types	12%	7.9%	8.9%	11.4%
	Adjusted Residual	2.2	-1.9	-1.1	
PhD	Count	13	0	4	17
	Expected Count	14.3	1.6	1.1	17
	% within Engagement Types	0.6%	0%	2.2%	0.6%
	Adjusted Residual	-0.9	-1.3	2.8	
Other	Count	72	9	8	89
	Expected Count	74.9	8.3	5.9	89
	% within Engagement Types	3.1%	3.5%	4.4%	3.3%
	Adjusted Residual	-0.9	0.3	0.9	
Total	Count	2302	254	180	2736
	% within Engagement Types	100%	100%	100%	100%

5.3.3 Proportions of HDI levels in the Engagement Types of Edraak

Overall, learners from countries with a medium HDI level seem to represent the highest proportion in all engagement types, see Figure -7. However, chi-square test showed **no significant** difference, in the proportions of countries with different HDI levels, between the engagement types ($X^2 = 10.830, df = 6, p > 0.05$), table 5.7.

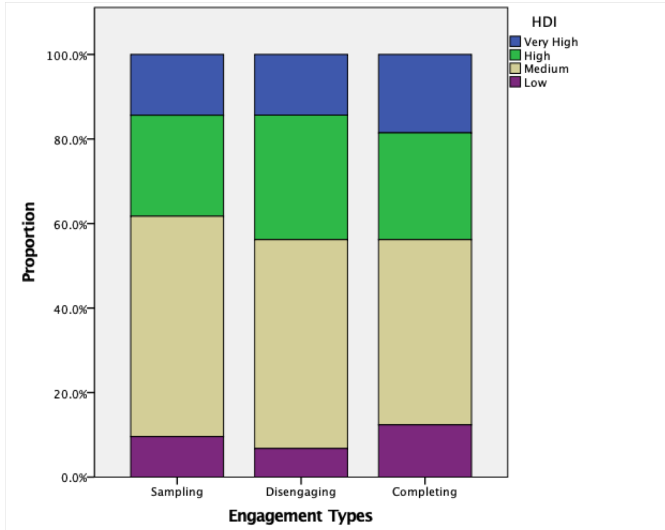


Figure 5-7: The proportions of HDI levels within the engagement types of Edraak learners

Table 5.7: Crosstabulation of the proportions of HDI level and the engagement types of Edraak learners

		Sampling	Disengaging	Completing	Total
Very High	Count	327	36	33	396
	Expected Count	333.1	36.8	26.1	396
	% within Engagement Types	14.4%	14.3%	18.5%	14.7%
	Adjusted Residual	-0.9	-0.1	1.5	
High	Count	543	74	45	662
	Expected Count	556.9	61.5	43.6	662
	% within Engagement Types	23.9%	29.5%	25.3%	24.5%
	Adjusted Residual	-1.7	1.9	0.3	
Medium	Count	1185	124	78	1387
	Expected Count	1166.8	128.8	91.4	1387
	% within Engagement Types	52.1%	49.4%	43.8%	51.3%
	Adjusted Residual	1.9	-0.6	-2.1	
Low	Count	218	17	22	257
	Expected Count	216.2	23.9	16.9	257
	% within Engagement Types	9.6%	6.8%	12.4%	9.5%
	Adjusted Residual	0.3	-1.6	1.3	
Total	Count	2273	251	178	2702
	% within Engagement Types	100%	100%	100%	100%

5.3.4 The Use of the Discussion Forum Among The Engagement Types of Edraak

In Section 4.4 we observed that the participation, as represented by the number of comments, of Edraak learners in the discussion forum was very low, which prevents us from performing any statistical analysis. Therefore, we can not explore the distribution of comments between the three engagement types of Edraak learners.

By looking at each engagement type, we can notice that “Completing” learners have a higher number of comments than the other two engagement types (29 comments for “Completing” learners, 3 comments for “Disengaging” learners and 7 comments for “Sampling” learners). This might indicate that communicating with other learners helped in completing the MOOC, which is consistent with the role of discussion in completing MOOCs [2].

5.3.5 Exposure to MOOC Materials (Watching Video Lectures) Between the Engagement Types of Edraak

Studying the differences between the engagement types of Edraak learners in their interaction with the MOOC materials, the ANOVA test showed a **significant** effect among the compared groups, $p < 0.001$. The Games-Howell post-hoc test showed that the mean number of video lectures watched by the “Sampling” learners was significantly lower than the other engagement types, see Figure -9 (for more information about the ANOVA and the post-hoc tests, see Section 3.3.3). This result was expected since Sampling learners dropped out early from the MOOC, as showed in Section 5.2 and Figure 5-2. Interestingly, video lecture watching between “Disengaging” and “Completing” learners were statistically similar, as “Completing” learners were expected to watch more video lectures.

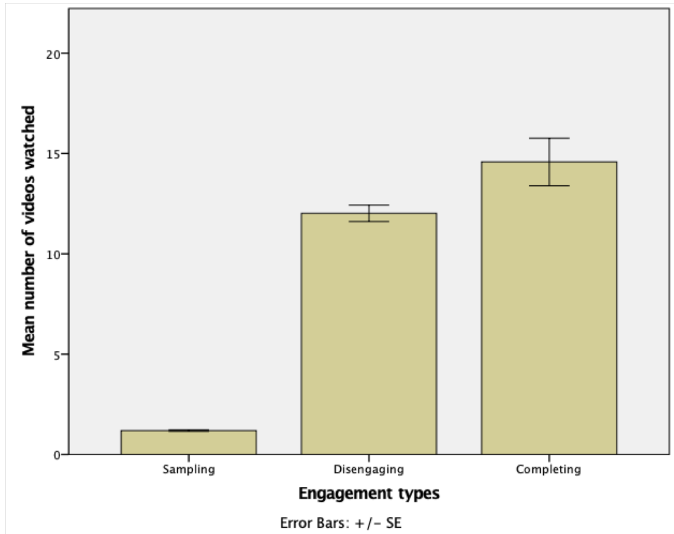


Figure 5-8: Engagement types interaction with video lectures

Remark. *In studying the differences among the Edraak's Engagement types, "Completing" learners had a more interaction in the discussion forum than "Disengaging" and "Sampling". At the opposite end, the "Sampling" learners had the lowest interaction with the video lectures.*

5.4 Discussion

Kizilcec et al. [1] studied the learners engagement with MOOCs on Coursera's platform for the purpose of identifying different learner engagement types, which enabled them to investigate the large dropout rate in MOOCs. Using a K-means clustering algorithm, the Kizilcec team managed to identify four engagement types (Sampling, Disengaging, Auditing and Completing), which are explained in detail in Section 2.2.3. Their clustering algorithm was based on computing an "engagement description" for each learner, which descries his/her weekly interaction with the MOOC contents, video lectures and assessment in the following way:

- When submitting the assessment on time, a learner was given a score of **3**
- When submitting the assessment late, a learner was given a score of **2**

- When watching videos only and not submitting the assessment, a learner was given a score of **1**
- For no engagement at all, a learner was given a score of **0**

Then, they used the L_1 norm on the engagement descriptions to calculate similarities among learners, and the result of the L_1 norm calculation was used to run the K-means clustering algorithm. In Section 2.2.2.2, we explained why we believe the L_1 norm step might have affected the clustering result, such that learners with different engagement descriptions may have had the same output.

In this chapter, we were interested in identifying the Arabic learners engagement types, which will be used in the nexy chapter to compare with Coursera's engagement types. This is to better understand the way Arabic learners engage with MOOCs and identify similarities and differences with English learners' engagement with MOOCs. This would help platform designers and MOOCs providers to improve their courses accordingly. Our data were generated from the Arabic MOOC platform "Edraak". For the purpose of this comparison, we sought to follow their clustering algorithm with the L_1 norm calculation. The only difference we made in our analysis was in the computation of the "engagement description", because the Edraak platform did not allow late submission, making our computation of the engagement description as follow:

- When submitting the assessment on time, a learner was given a score of 3
- When watching videos only and not submitting the assessment, a learner was given a score of 2
- For no engagement at all, a learner was given a score of 1

The first week of the Edraak MOOC was an introduction to the course and had only one video that welcomes the learners and guides them to download a programming platform with no assessment. This would affect the application of a scoring system equally on all weeks, therefore, it was better to exclude the first week data from the engagement description calculation. The clustering of the Edraak data was performed using the K-means clustering algorithm with K= 3, which resulted in three engagement types: Sampling, Disengaging and Completing, see Section 5.2.

As an attempt to find the potential reasons behind the obtained engagement types, we analyzed the learner demographics with statistical tests. Our results showed that "Completing" learners had more interaction in the discussion forum than "Disengaging" and "Sampling", which might link the communication among learners to completing the MOOC. This finding was consistent with a study done by Ferguson and Clow [2], who linked a high communicating rate with completing a MOOC.

Another finding related learners' interaction with the video lectures showed that, as expected, "Sampling" learners had the lowest interaction as they dropped out early in the MOOC. However, interestingly, there was no difference in the video lecture interaction between "Completing" and "Disengaging" learners, despite the fact that Disengaging learners had dropped out in the middle of the MOOC, whereas Completing learners had completed the MOOC. This clearly showed that Completing learners were interested in only completing the weekly assessment which allowed them to complete the MOOC and allowed them to obtain a completion certificate from the MOOC platform without concern about the MOOC content. Table 5.8 summarise the results of the statistical analyses in this chapter.

In conclusion, using the Kizilcec et al.'s clustering algorithm [1], we managed to identify three engagement types that represented Edraaks learners. Our analyses of these engagement types raised some more questions: For the "Sampling" learners, what caused their quick dropout? And what was the objective of their enrolment? For the "Disengaging" learners, why did they return to the course after the first week unlike Sampling? And what was the aim of their enrolment? For "Completing" learners, what was the motivation behind completing the course? Were they interested in requesting a completion certificate? If so, why? These questions could be addressed and answered by performing further qualitative studies.

Table 5.8: Summary of the statistical analyses

Questions	Statistical test	Significance $p < 0.05$	Strength of significance (Cramer's V)
1. Is there a significant difference in the gender proportion among the engagement types? If so, what is the strength of this significant?	Chi-Square	✔	Small
2. Is there a significant difference in the learners' qualifications among the engagement types? If so, what is the strength of this significant?	Chi-Square	✗	-
3. Is there a significant difference in the distribution over countries with different HDI levels among the engagement types? If so, what is the strength of this significant?	Chi-Square	✗	-
4.Are there significant differences in the exposure to the course materials among the engagement types?	ANOVA	✔	-

Chapter 6

Comparison Between Edraak's and Coursera's learners

In Chapter 5, we analysed Edraak's learner data, using the Kizilcec et al. [1] method to investigate their engagement with MOOCs. We found that three out of the four engagement ttypes identified by Kizilcec et al. demonstrated by Edraak learners. These three engagement types were: Sampling, Disengaging and Completing.

In this chapter, we compare the Kizilcec et al. engagement types, which represent Coursera's learners, with our engagement types, which represent Edraak's learners. This is to answer our third research question (**How is the engagement of Arabic-speaking learners in a MOOC similar to or different from the engagement of English-speaking learners in a MOOC?**). Table 6.1 shows learner data from the Coursera and Edraak platforms, where the Coursera's column contains data that has been taken from Kizilcec et al. [1].

Table 6.1: Coursera VS Edraak

		Coursera	Edraak
No. of Learners		94091	2736
Gender	Males	76 (%)	68 (%)
	Females	24 (%)	32 (%)
HDI level	Very High	65 (%)	14 (%)
	High	14 (%)	24 (%)
	Medium	18 (%)	51 (%)
	Low	3 (%)	9 (%)
Engagement Types	Auditing	7 (%)	-
	Completing	17 (%)	7 (%)
	Disengaging	18 (%)	9 (%)
	Sampling	58 (%)	84 (%)

To compare the engagement types of Coursera and Edraak, we applied the following statistical approach:

1. Is there a significant difference in the engagement types proportion between Edraak's and Coursera's learners?. If so, what is the strength of this significant?
2. Is there a significant difference in the gender proportion between Edraak's and Coursera's learners? If so, what is the strength of this significant?
   2.1. Is there a significant difference in the gender proportion between Edraak's and Coursera's Engagement types? If so, what is the strength of this significant?
3. Is there a significant difference in the distribution over countries with different HDI levels between Edraak's and Coursera's learners? If so, what is the strength of this significant?
   3.1. Is there a significant difference in the distribution over countries with different HDI levels between Edraak's and Coursera's Engagement types? If so, what is the strength of this significant?
4. Is there a significant difference in the Use of Discussion Forums between Edraak's and Coursera's learners? If so, what is the strength of this significant?

To measured these significances, this approach used the **Chi-square** test with significance level of $p < 0.05$ and adjusted residual value of ± 1.96. Then a Cramer's V was used to determine the strength of these significances (for more information about Chi-Square, adjusted residual and Cramer's V, see Sections 3.3.1, 3.3.1.1 and 3.3.1.2).

6.1 Proportions of Engagement Types

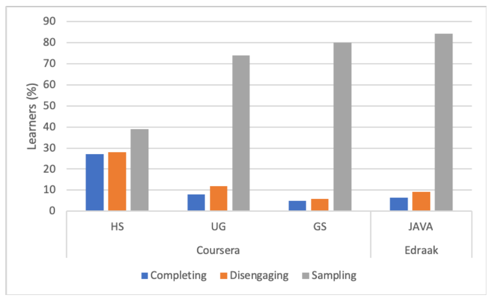


Figure 6-1: Engagement Types proportions between Coursera and Edraak

To explore the learner engagement of Edraak, we compared the obtained engagement types, in Section 5.2, with Coursera's engagement types, in Section 2.2.3. Coursera and Edraak shared three engagement types: Completing, Disengaging and Sampling. Looking at the proportions of these engagement types, see Figure 6-1, we noticed that the majority of learners from both platform were Sampling, and Edraak has the highest percentage of learners falling in this engagement type. We can also notice that the proportions of the three engagement types in the JAVA MOOC on Edraak were similar to the undergraduate-level (UG) and the graduate-level (GS) MOOCs, and very different from the high-school level (HS) MOOC on Coursera.

Table 6.2: Engagement Types * MOOCs Crosstabulation: Coursera's HS MOOC and Edraak's JAVA MOOC

		Coursera HS	Edraak JAVA	Total
Auditing	Count	2766	0	2766
	Expected Count	2611	155	2766
	% within platform	6%	0%	5.7%
	Adjusted Residual	13.2	-13.2	
Completing	Count	12446	180	12626
	Expected Count	11918.6	707.4	12626
	% within platform	27%	6.6%	25.9%
	Adjusted Residual	23.7	-23.7	
Disengaging	Count	12907	254	13161
	Expected Count	12423.6	737.4	13161
	% within platform	28%	9.3%	27%
	Adjusted Residual	21.4	-21.4	
Sampling	Count	17977	2302	20279
	Expected Count	19142.8	1136.2	20279
	% within platform	39%	84.1%	41.5%
	Adjusted Residual	-46.6	46.6	
Total	Count	46096	2736	48832
	% within platform	100%	100%	100%

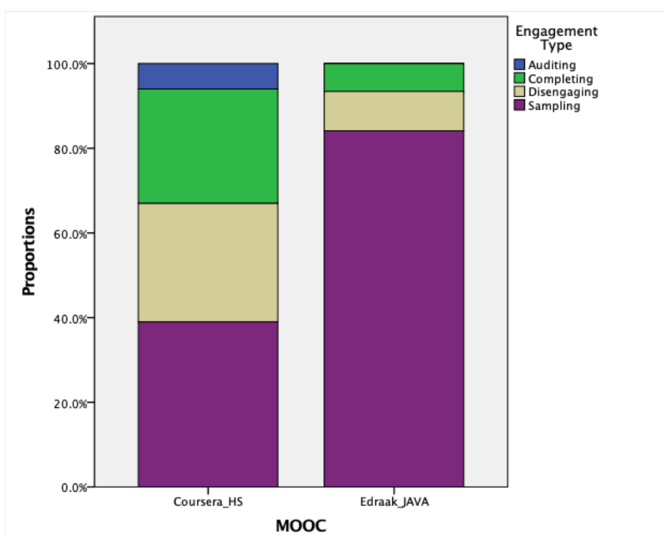


Figure 6-2: The proportions of the engagement types of learners in Coursera's HS and Edraak's

When comparing the three engagement types between the two platforms (see Table 6.2 and Figure 6-2), Coursera's HS MOOC had more "Completing" and "Disengaging" learners and less "Sampling" than expected. On the contrary, Edraak's "Completing" and "Disengaging" learners were far less than expected, whereas "Sampling" learners were more. A chi-square test showed a "**significant**" difference among the engagement types of the platforms (Coursera and Edraak) (X^2 = 2183.57, *df* = 3, $p < 0.001$). However, measuring the strength of this significance using Cramer's V test showed that was "**Medium**", with a value of 0.21, which is considered acceptable [116]. This could mean that the learners engagement with the MOOC was influenced by the platform.

Table 6.3: Engagement Types * MOOCs Crosstabulation: Coursera's UG MOOC and Edraak's JAVA MOOC

		Coursera UG	Edraak JAVA	Total
Auditing	Count	1613	0	1613
	Expected Count	1464	149	1613
	% within platform	6%	0%	5.4%
	Adjusted Residual	13.2	-13.2	
Completing	Count	2151	180	2331
	Expected Count	2115.7	215.3	2331
	% within platform	8%	6.6%	7.9%
	Adjusted Residual	2.6	-2.6	
Disengaging	Count	3226	254	3480
	Expected Count	3158.6	321.4	3480
	% within platform	12%	9.3%	11.7%
	Adjusted Residual	4.2	-4.2	
Sampling	Count	19897	2302	22199
	Expected Count	20148.7	2050.3	22199
	% within platform	74%	84.1%	74.9%
	Adjusted Residual	-11.7	11.7	
Total	Count	26887	2736	29623
	% within platform	100%	100%	100%

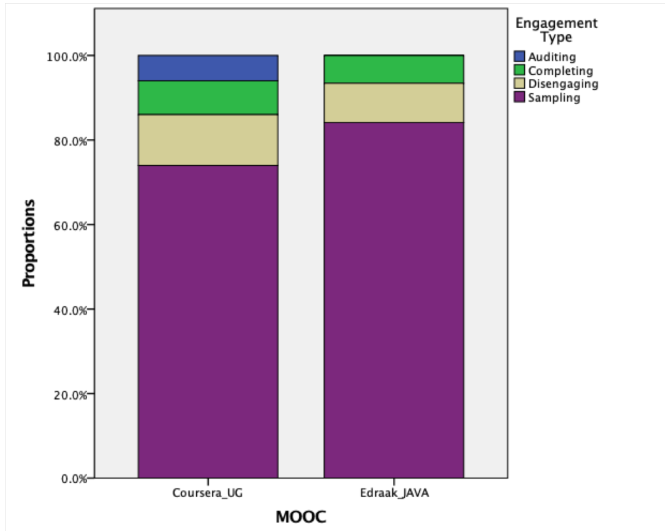


Figure 6-3: The proportions of the engagement types of learners in Coursera's UG and Edraak's

Next, we looked at the second MOOC on Coursera (UG MOOC) (see Table 6.3 and Figure 6-3), and compared it to Edraak's JAVA MOOC. The data showed that Coursera's UG MOOC had more "Completing" and "Disengaging" learners and less "Sampling" than expected. In comparison, less "Completing" and "Disengaging" learners and more "Sampling" than expected were found in Edraak's JAVA MOOC. Analysing the data with the chi-square test showed a significant difference among the engagement types of the platform used (Coursera and Edraak) ($X^2 = 220.13$, $df = 3$, $p < 0.001$). However, Cramer's V test showed a "**Small**" strength of significance, with a value of 0.09, which is considered **unacceptable** [116]. This tells us that the observed significance might resulted from the high difference in the sample size. Therefore, we conclude that, within Coursera's UG and Edraak's JAVA MOOCs, the platform is less considered to have an influence in the learners engagement.

Table 6.4: Engagement Types * MOOCs Crosstabulation: Coursera's GS MOOC and Edraak's JAVA MOOC

		Coursera GS	Edraak JAVA	Total
Auditing	Count	1900	0	1900
	Expected Count	1682	218	1900
	% within platform	9%	0%	8%
	Adjusted Residual	16.4	-16.4	
Completing	Count	1055	180	1235
	Expected Count	1093.3	141.7	1235
	% within platform	5%	6.6%	5.2%
	Adjusted Residual	-3.5	3.5	
Disengaging	Count	1267	254	1521
	Expected Count	1346.5	174.5	1521
	% within platform	6%	9.3%	6.4%
	Adjusted Residual	-6.6	6.6	
Sampling	Count	16886	2302	19188
	Expected Count	16986.3	2201.7	19188
	% within platform	80%	84.1%	80.5%
	Adjusted Residual	-5.1	5.1	
Total	Count	21108	2736	23844
	% within platform	100%	100%	100%

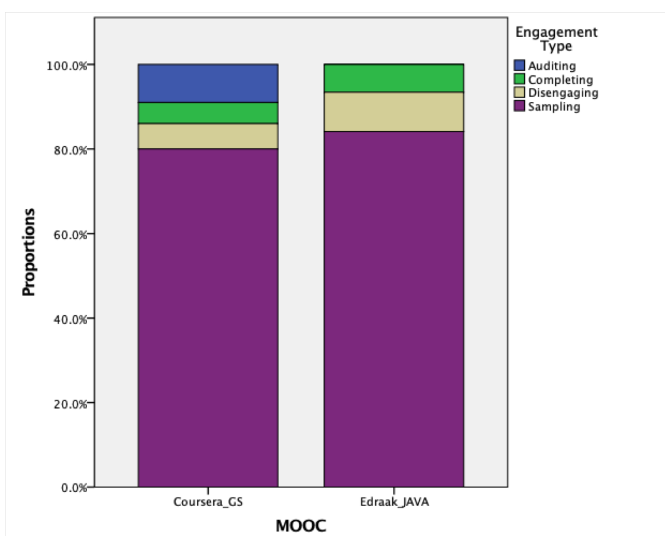


Figure 6-4: The proportions of the engagement types of learners in Coursera's GS and Edraak's

Finally, we compared Coursera's GS MOOC and Edraak's JAVA MOOC, see Table 6.4 and Figure 6-4. The chi-squared test produced a similar result to Coursera's HS and UG MOOC, where it showed a significant difference between the engagement types and the platforms ($X^2 = 303.997$, $df = 3$, $p < 0.001$). Testing the strength of this significant result using Cramer's V test showed that with a value of 0.11, there was a "**Small**" strength, which is considered **unacceptable** [116]. This presumably resulted from the high difference in the sample size between the two platforms.

Remark. *Among the studied Cureara MOOCs, the HS and UG MOOCs had a higher "Completing" and "Disengaging" rate than Edraak's JAVA MOOC, whereas the same engagement types in Coursera's GS MOOC appeared at a lower rate than Edraak's JAVA MOOC. Moreover, Edraak's JAVA MOOC had the highest "Sampling" rate over all of the Coursear MOOCs. Overall, the number of learners who fell into the "Sampling" engagement type was the highest among all observed types in both platforms.*

6.2 Gender Proportion

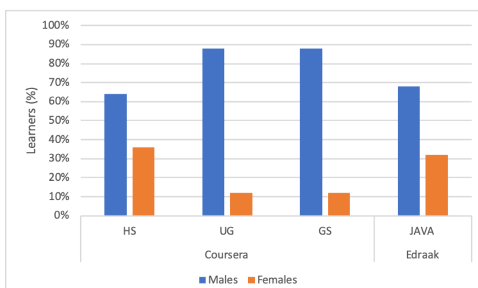


Figure 6-5: Gender proportions of Coursera's three MOOCs and Edraak's MOOC

Looking at gender proportions of learners from Coursera and Edraak (see Figure 6-5), we can say that in both platforms, the enrolment of male learners were more than female learners. The gender proportions of Edraak's JAVA MOOC was very similar to the high-school level MOOC (HS) on Coursera. To test if there is a significant different in gender proportion between the two platforms, we ran a chi-squared test between each of the three MOOCs from Coursera and our MOOC from Edraak.

Table 6.5 shows the crosstabulation of gender and platforms between Coursera's

HS MOOC and Edraak's Java MOOC. A significant difference was found between the two platforms within each gender (X^2 = 18.649, df = 1 and $p < 0.001$). Similar outcome was shown when comparing Edraak's JAVA MOOC with Coursera's UG and GS MOOCs, Table 6.6 (X^2 = 823.103, df = 1 and $p < 0.001$) and Table 6.7 (X^2 = 781.163, df = 1 and $p < 0.001$), respectively. Using Cramer's V test to measure the strength of these significant differences showed a "**Small**" strength of significant between gender and platforms in all three tests, with values of 0.02, 0.17 and 0.18, respectively, which is considered **unacceptable** [116].

Table 6.5: Gender * MOOCs Crosstabulation: Coursera's HS MOOC and Edraak's Java MOOC

		Coursera HS	Edraak	Total
Female	Count	16595	872	17467
	Expected Count	16489.4	977.6	17467
	% within platform	36%	31.9%	35.8%
	Adjusted Residual	4.3	-4.3	
Male	Count	29501	1861	31362
	Expected Count	29606.6	1755.4	31362
	% within platform	64%	68.1%	64.2%
	Adjusted Residual	-4.3	4.3	
Total	Count	46096	2733	48829
	% within platform	100%	100%	100%

Table 6.6: Gender * MOOCs Crosstabulation: Coursera's UG Course and Edraak's Java Course

		Coursera UG	Edraak	Total
Female	Count	3226	872	4098
	Expected Count	3719.9	378.1	4098
	% within platform	12.0%	31.9%	13.8%
	Adjusted Residual	-28.7	28.7	
Male	Count	23661	1861	25522
	Expected Count	23167.1	2354.9	25522
	% within platform	88.0%	68.1%	86.2%
	Adjusted Residual	28.7	-28.7	
Total	Count	26887	2733	29620
	% within platform	100%	100%	100%

Table 6.7: Gender * MOOCs Crosstabulation: Coursera's GS Course and Edraak's Java Course

		Coursera GS	Edraak	Total
Female	Count	2534	872	3406
	Expected Count	3015.6	390.4	3406
	% within platform	12%	31.9%	14.3%
	Adjusted Residual	-28	28	
Male	Count	18576	1861	20437
	Expected Count	18094.4	2342.6	20437
	% within platform	88%	68.1%	85.7%
	Adjusted Residual	28	-28	
Total	Count	21110	2733	23843
	% within platform	100%	100%	100%

Next, we tested whether the gender proportions are different between the three engagement types of Coursera and Edraak. We used Chi-Square test to compare gender proportions between: Coursera's Completing and Edraak's Completing; Coursera's Disengaging and Edraak's Disengaging; and Coursera's Sampling and Edraak's Sampling.

Starting with testing the gender proportions difference between Edraak's JAVA MOOC and Coursera's HS MOOC, see Table 6.8. With the chi-squared results, $X^2 = 1243.031$, $df = 3$ and $p < 0.001$ for female proportions, and $X^2 = 1110.962$, $df = 3$ and $p < 0.001$ for male proportions, there was significant difference in gender proportions between Corsera and Edraak within ALL engagement types (Completing, Disengaging, and Sampling). Testing the strength of the significant for the male data using Cramer's V test showed a "**Small**" strength of significant with value of 0.18, which is considered **unacceptable** [116]. Also, a similar strength of significant was found between the female proportions within each engagement type and platform with value 0.28. The high difference in the sample size might be the cause of the small strength of significant obtained from gender proportions in each platform.

Table 6.8: Gender * MOOCs Crosstabulation: Coursera's HS Course and Edraak's Java Course

			Coursera HS	Edraak	Total
Female	Auditing	Count	892	0	892
		Expected Count	842.2	49.8	892
		% within platform	6%	0%	5.7%
		Adjusted Residual	7.5	-7.5	
	Completing	Count	49610	42	4652
		Expected Count	4392.3	259.7	4652
		% within platform	31.3%	4.8%	29.8%
		Adjusted Residual	16.6	-16.6	
	Disengaging	Count	4964	77	5041
		Expected Count	4759.5	281.5	5041
		% within platform	33.7%	8.8%	32.3%
		Adjusted Residual	15.2	-15.2	
	Sampling	Count	4280	753	5033
		Expected Count	4752	281	5033
		% within platform	29%	86.4%	32.2%
		Adjusted Residual	-35.2	35.2	
	Total	Count	14746	872	15618
		% within platform	100%	100%	100%
Male	Auditing	Count	1874	0	1874
		Expected Count	1769	105	1874
		% within platform	6%	0%	5.6%
		Adjusted Residual	10.9	-10.9	
	Completing	Count	7836	138	8120
		Expected Count	7527.2	446.8	8120
		% within platform	25%	7.4%	24.4%
		Adjusted Residual	17.3	-17.3	
	Disengaging	Count	7943	177	8120
		Expected Count	7665	455	8120
		% within platform	25.3%	9.5%	24.4%
		Adjusted Residual	15.4	-15.4	
	Sampling	Count	13697	1546	15243
		Expected Count	14388.8	854.2	15243
		% within platform	43.7%	83.1%	45.9%
		Adjusted Residual	-33.1	33.1	
	Total	Count	31350	1861	33211
		% within platform	100%	100%	100%

We then repeated the same analysis on Coursera's UG MOOC to compare its gender proportions with Edraak's JAVA MOOC, see Table 6.9. With chi-square results of $X^2 = 47.862$, $df = 3$ and $p < 0.001$, there was significant differences in the female proportions between Corsera and Edraak within the "Disengaging" and "Sampling" engagement types (adjusted residual values of ±3.8 and ±5.7, respectively). Similarly, there was significant differences in the male proportions between Corsera and Edraak within the "Disengaging" and "Sampling" engagement types with a chi-square results of $X^2 = 153.742$, $df = 3$ and $p < 0.001$ (adjusted residual values of ± 2.8 and ± 9.3, respectively). Cramer's V test showed that the strength of significant between the gender proportions within each engagement type and the platform is "**Small**" with values of 0.09 and 0.07 for female and male proportions, respectively, which is considered **unacceptable** [116].

Table 6.9: Gender * MOOCs Crosstabulation: Coursera's UG Course and Edraak's Java Course

			Coursera UG	Edraak	Total
Female	Auditing	Count	141	0	141
		Expected Count	117.7	23.3	141
		% within platform	3.2%	0%	2.7%
		Adjusted Residual	5.4	-5.4	
	Completing	Count	242	42	284
		Expected Count	237.1	46.9	284
		% within platform	5.5%	4.8%	5.4%
		Adjusted Residual	0.8	-0.8	
	Disengaging	Count	597	77	674
		Expected Count	562.8	111.2	674
		% within platform	13.5%	8.8%	12.8%
		Adjusted Residual	3.8	-3.8	
	Sampling	Count	3431	753	4184
		Expected Count	3493.4	690.6	4184
		% within platform	77.8%	86.4%	79.2%
		Adjusted Residual	-5.7	5.7	
	Total	Count	4411	872	5283
		% within platform	100%	100%	100%
Male	Auditing	Count	1472	0	1472
		Expected Count	1359.4	112.6	1472
		% within platform	6.5%	0%	6%
		Adjusted Residual	11.4	-11.4	
	Completing	Count	1909	138	2047
		Expected Count	1890.5	156.5	2047
		% within platform	8.5%	7.4%	8.4%
		Adjusted Residual	1.6	-1.6	
	Disengaging	Count	2629	177	2806
		Expected Count	2591.4	214.6	2806
		% within platform	11.7%	9.5%	11.5%
		Adjusted Residual	2.8	-2.8	
	Sampling	Count	16466	1546	18012
		Expected Count	16634.7	1377.3	18012
		% within platform	73.3%	83.1%	74%
		Adjusted Residual	-9.3	9.3	
	Total	Count	22476	1861	24337
		% within platform	100%	100%	100%

Finally, we ran the same analysis on Coursera's GS MOOC to compare its gender proportions with Edraak's, see Table 6.10. We found that there was a significant difference in the female proportions between Corsera and Edraak within the "Disengaging" engagement type only, with a chi-square result of $X^2 = 71.713$, $df = 3$ and $p < 0.001$ (adjusted residual of ± 7.3). On the other hand, there were significant differences in the male proportions between Corsera and Edraak within ALL engagement types, with a chi-square result of $X^2 = 223.694$, $df = 3$ and $p < 0.001$. Again, Cramer's V test showed that the strength of significant between the gender proportions within each engagement type and the platform was "**Small**" with values of 0.15 and 0.1 for female and male proportions, respectively, which is considered **unacceptable** [116]

Table 6.10: Gender * MOOCs Crosstabulation: Coursera's GS Course and Edraak's Java Course

			Coursera GS	Edraak	Total
Female	Auditing	Count	167	0	167
		Expected Count	124.3	42.7	167
		% within platform	6.6%	0%	4.9%
		Adjusted Residual	7.8	-7.8	
	Completing	Count	101	42	143
		Expected Count	106.5	36.5	143
		% within platform	4%	4.8%	4.2%
		Adjusted Residual	-1.1	1.1	
	Disengaging	Count	135	77	212
		Expected Count	157.8	54.2	212
		% within platform	5.3%	8.8%	6.2%
		Adjusted Residual	-3.7	3.7	
	Sampling	Count	2137	753	2890
		Expected Count	2151.4	738.6	2890
		% within platform	84.1%	86.4%	84.7%
		Adjusted Residual	-1.6	1.6	
	Total	Count	2540	872	3412
		% within platform	100%	100%	100%
Male	Auditing	Count	1733	0	1733
		Expected Count	1575.1	157.9	1733
		% within platform	9.3%	0%	8.5%
		Adjusted Residual	13.8	-13.8	
	Completing	Count	954	138	1092
		Expected Count	992.5	99.5	1092
		% within platform	5.1%	7.4%	5.3%
		Adjusted Residual	-4.2	4.2	
	Disengaging	Count	1132	177	1309
		Expected Count	1189.8	119.2	1309
		% within platform	6.1%	9.5%	6.4%
		Adjusted Residual	-5.7	5.7	
	Sampling	Count	14749	1546	16295
		Expected Count	14810.6	1484.4	16295
		% within platform	79.4%	83.1%	79.8%
		Adjusted Residual	-3.7	3.7	
	Total	Count	18568	1861	20429
		% within platform	100%	100%	100%

Remark. *Overall, There was more male enrolment than female in both Coursera and Edraak. The enrolment of male learners was significantly higher in Coursera than Edraak, whereas the enrolment of female learners was significantly higher in Edraak than Coursera. Although the gender proportions were significantly different between Coursera's three MOOCs and Edraak's MOOC, the association between gender and platform was not strong.*

Although, significant differences were found in gender proportions within different engagement types between Coursera's three MOOCs and Edraak's MOOC, the strength of these significant differences were small. This might have been due to the sample size.

6.3 HDI Levels Proportion

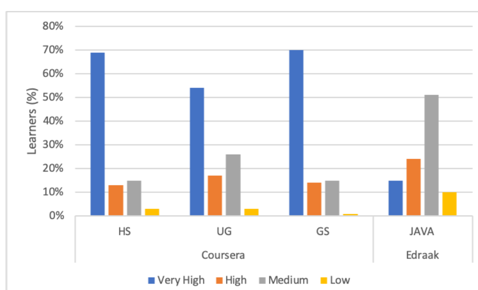


Figure 6-6: HDI level proportions between Coursera and Edraak

HDI is a well-known index to classify a country's development using three parameters: life expectancy, education, and income [98] (see Appendix A). To follow our analysis of the HDI of Edraak's learners in Section 5.2, we compare it with Coursera's learner's HDI [1]. Looking at the proportion of learners from different HDI level countries from Coursera and Edraak, see Figure 6-6, we noticed some major differences. The majority of Coursera's learners were from countries that are Very High in the development index (50% to 70%), whereas for Edraak, the majority of learners were from countries that are Medium in the development index (51%).

When comparing Edraak's JAVA MOOC with all Coursera's MOOCs (Tables 6.11, 6.12 and 6.13), we found that the differences in all HDI levels proportions between the learners of these MOOCs were "**significant**", with a chi-squared result of $X^2 =$

3692.448, $df = 3$ and $p < 0.001$ for Coursera's HS MOOC, $X^2 = 1678.450$, $df = 3$ and $p < 0.001$ for Coursera's UG MOOC, and $X^2 = 3861.975$, $df = 3$ and $p < 0.001$ for Coursera's GS MOOC. Cramer's V test of strength of significance showed a "**Medium**" strength of significance between the HDI levels and platforms when comparing Edraak's MOOC with Coursera's HS and UG MOOCs, with values of 0.28 and 0.24, respectively. Whereas, a "**Large**" strength of significance between the HDI levels and platforms was found when comparing Edraak's MOOC with the GS MOOC of Coursera with value = 0.4. These strengths are considered acceptable [116]. This outcome might imply that MOOCs in less developed Arab countries are more considered as a source of education than western countries.

Table 6.11: HDI level * MOOCs Crosstabulation: Coursera's HS Course and Edraak's Java Course

		Coursera HS	Edraak	Total
Very High	Count	31806	396	32202
	Expected Count	30418.9	1783.1	32202
	Adjusted Residual	58	-58	
High	Count	5993	662	6655
	Expected Count	6286.5	368.5	6655
	Adjusted Residual	-16.9	16.9	
Medium	Count	6914	1387	8301
	Expected Count	7841.4	459.6	8301
	Adjusted Residual	-48.9	48.9	
Low	Count	1383	257	1640
	Expected Count	1549.2	90.8	1640.0
	Adjusted Residual	-18.3	18.3	
Total	Count	46096	2702	48798

Table 6.12: HDI level * MOOCs Crosstabulation: Coursera's UG Course and Edraak's Java Course

		Coursera UG	Edraak	Total
Very High	Count	14519	396	14915
	Expected Count	13553	1362	14915
	Adjusted Residual	39	-39	
High	Count	4571	662	5233
	Expected Count	4755.1	477.9	5233
	Adjusted Residual	-9.7	9.7	
Medium	Count	6991	1387	8378
	Expected Count	7612.9	765.1	8378
	Adjusted Residual	-27.9	27.9	
Low	Count	806	257	1063
	Expected Count	965.9	97.1	1063
	Adjusted Residual	-17.3	17.3	
Total	Count	26887	2702	29589

Table 6.13: HDI level * MOOCs Crosstabulation: Coursera's GS Course and Edraak's Java Course

		Coursera GS	Edraak	Total
Very High	Count	14776	396	15172
	Expected Count	13450.3	1721.7	15172
	Adjusted Residual	56.3	-56.3	
High	Count	2955	662	3617
	Expected Count	3206.5	410.5	3617
	Adjusted Residual	-14.3	14.3	
Medium	Count	3166	1387	4553
	Expected Count	4036.3	516.7	4553
	Adjusted Residual	-45.2	45.2	
Low	Count	211	257	468
	Expected Count	414.9	53.1	468
	Adjusted Residual	-30.0	30.0	
Total	Count	21108	2702	23810

Kizilcec et al. published the HDI level of their engagement types for Coursera's GS MOOC only [1]. Since our earlier result, in Table 6.13, indicates a strong relation between the HDI levels and platforms, between Coursera's GS MOOC and Edraak's MOOC, we sought to examine the HDI levels and the engagement types from Coursera's GS MOOC with Edraaks' to identify if a similar relation would be found. Figure 6-7 shows the HDI level proportions for the engagement types of the two MOOCs, Coursera and Edraak, that we studied.

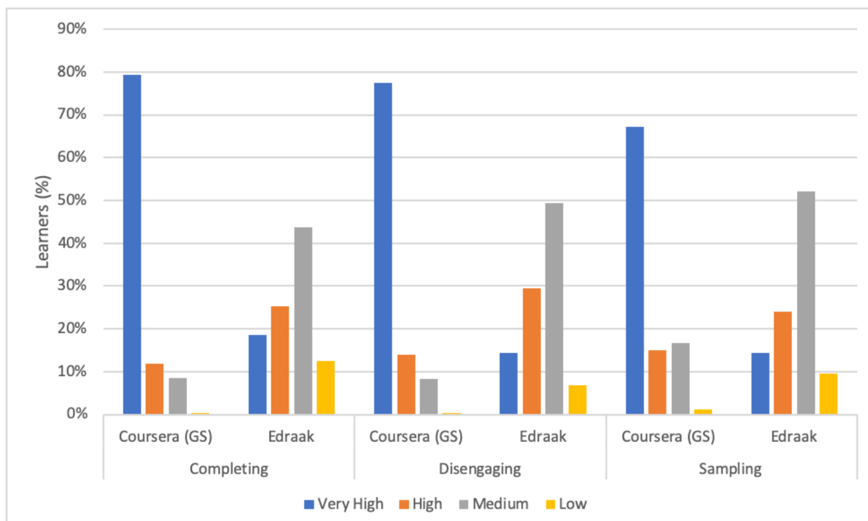


Figure 6-7: HDI level proportions between the engagement types of Coursera and Edraak

The data in Table 6.14 showed a significant difference between Coursera's and Edraak's learners proportions in the "Sampling" engagement type in the Very High HDI level, with a chi-square result of $X^2 = 62.112$, $df = 3$ and $p < 0.001$, and in the High HDI level, with a chi-square result of $X^2 = 57.919$, $df = 3$ and $p < 0.001$. However, both the Very High and High HDI levels showed a "**Small**" strength of significant between the two platforms and the HDI levels with Cramer's V test values of 0.06 and 0.13, respectively, which are considered **unacceptable** [116]. This result might occurred due to the difference in the sample size.

All engagement types within the Medium HDI level showed a "**significant**" difference between Coursera's and Edraak's learners with a chi-square result of $X^2 = 185.798$, $df = 3$ and $p < 0.001$, and a Cramer's V test value of 0.2 that indicates a "**Medium**" strength of significance between the two platforms and the HDI levels. Similarly, a "**significant**" difference between Coursera's and Edraak's learners proportions in the "Completing" engagement type was found in the Low HDI level with a chi-square result of $X^2 = 47.847$, $df = 3$ and $p < 0.001$, but with a "**Large**" strength of significance between the two platforms and the HDI levels with a Cramer's V test value of 0.32. Both differences are considered acceptable [116]. Interestingly, this analysis could indicate that learners from Low HDI countries were more likely to complete MOOCs in Edraak platform.

Table 6.14: HDI level * MOOCs Crosstabulation: Coursera's GS Course and Edraak's Java Course

			Coursera GS	Edraak	Total
Very High	Auditing	Count	1921	0	1921
		Expected Count	1870.9	50.1	1921
		% within platform	13.0%	0%	12.7%
		Adjusted Residual	7.7	-7.7	
	Completing	Count	1182	33	1215
		Expected Count	1183.3	31.7	1215
		% within platform	8%	8.3%	8%
		Adjusted Residual	-0.2	0.2	
	Disengaging	Count	1478	36	1514
		Expected Count	1474.5	39.5	1514
		% within platform	10.0%	9.1%	10%
		Adjusted Residual	0.6	-0.6	
	Sampling	Count	10195	327	10522
		Expected Count	10247.4	274.6	10522
		% within platform	69%	82.6%	69.4%
		Adjusted Residual	-5.8	5.8	
	Total	Count	14776	396	15172
		% within platform	100%	100%	100%

Continued on next page

Table 6.14 – continued from previous page

			Coursera GS	Edraak	Total
High	Auditing	Count	236	0	236
		Expected Count	192.8	43.2	236
		% within platform	8%	0%	6.5%
		Adjusted Residual	7.5	-7.5	
	Completing	Count	177	45	222
		Expected Count	181.4	40.6	222
		% within platform	6%	6.8%	6.1%
		Adjusted Residual	-0.8	0.8	
	Disengaging	Count	266	74	340
		Expected Count	277.8	62.2	340
		% within platform	9%	11.2%	9.4%
		Adjusted Residual	-1.7	1.7	
	Sampling	Count	2275	543	2818
		Expected Count	2302.1	515.9	2818
		% within platform	77%	82%	77.9%
		Adjusted Residual	-2.8	2.8	
	Total	Count	2954	662	3616
		% within platform	100%	100%	100%
Medium	Auditing	Count	348	0	348
		Expected Count	242	106	348
		% within platform	11%	0%	7.6%
		Adjusted Residual	12.8	-12.8	
	Completing	Count	127	78	205
		Expected Count	142.5	62.5	205
		% within platform	4%	5.6%	4.5%
		Adjusted Residual	-2.4	2.4	
	Disengaging	Count	158	124	282
		Expected Count	196.1	85.9	282
		% within platform	5.0%	8.9%	6.2%
		Adjusted Residual	-5.1	5.1	
	Sampling	Count	2533	1185	3718
		Expected Count	2585.4	1132.6	3718
		% within platform	80%	85.4%	81.7%
		Adjusted Residual	-4.4	4.4	
	Total	Count	3166	1387	4553
		% within platform	100%	100%	100%

Continued on next page

Table 6.14 – continued from previous page

			Coursera GS	Edraak	Total
Low	Auditing	Count	30	0	30
		Expected Count	13.5	16.5	30
		% within platform	14.2%	0%	6.4%
		Adjusted Residual	6.2	-6.2	
	Completing	Count	4	22	26
		Expected Count	11.7	14.3	26
		% within platform	1.9%	8.6%	5.6%
		Adjusted Residual	-3.1	3.1	
	Disengaging	Count	8	17	25
		Expected Count	11.3	13.7	25
		% within platform	3.8%	6.6%	5.3%
		Adjusted Residual	-1.4	1.4	
	Sampling	Count	169	218	387
		Expected Count	174.5	212.5	387
		% within platform	80.1%	84.8%	82.7%
		Adjusted Residual	-1.3	1.3	
	Total	Count	211	257	468
		% within platform	100%	100%	100%

Remark. *The proportions of HDI levels were significantly different between Coursera's three MOOCs and Edraak's MOOC. The vast majority of learners from Coursera's three MOOCs were from countries with Very High HDI levels, and very few of them were from Low HDI level countries. This was significantly different from Edraak, where the proportions of learners from Low and Medium HDI countries were higher than Coursera's. In addition, Edraak's learners from Low HDI countries seemed to complete MOOCs more than Coursera's learners from countries with the same HDI level.*

6.4 Use of Discussion Forum

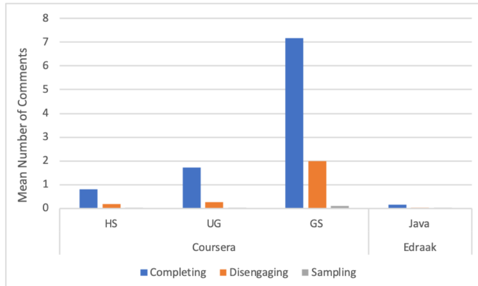


Figure 6-8: Learners Use of Discussion Forum in Coursera and Edraak

Table 6.15: Use of Discussion Forum between Coursera's and Edraak's Learners: Mean Number of Comments

		Completing	Disengaging	Sampling
Coursera	HS	0.79	0.19	0.02
	UG	1.71	0.25	0.02
	GS	7.18	1.98	0.09
Edraak	JAVA	0.16	0.01	0.003

As we mentioned before, in Section 4.4, the number of comments of Edraak learners in the discussion forum is very low. Therefore, we can not use any statistical analysis to compare the use of discussion forum within each engagement type of Coursera's and Edraak's MOOCs.

Figure 6-8 and Table 6.15 show the mean number of comments made by learners in each engagement type from Coursera's three MOOCs and Edraak's. It is clear that learners from Coursera's GS MOOC had the highest engagement with the discussion forum. Looking at the mean numbers of comments, see Table 6.15, we can see that in all the MOOCs, Completing learners had the highest engagement with the discussion forums.

6.5 Discussion

MOOCs are known as useful open educational resources that have grown rapidly in the Western world. Their popularity is recently gaining more attention in the Arabic world [62]. This chapter is an extension of our analysis of Edraak's engagement types in Chapter 5, and aimed to compare them with Coursera's engagement types to explore any similarities or differences in the engagement patterns that can benefit MOOC designers to better understand these patterns of the Arabic platform users and accordingly to design suitable MOOCs for them.

Kizilcec et al. [1] studied the learners patterns of interaction with video lectures and assessments of three MOOCs from the Coursera platforms (HS, UG and GS). Their study resulted in four engagement types; Completing, Auditing, Disengaging and Sampling. To be able to compare the engagement patterns of Edraak with Coursera, we applied their clustering method on Edraak's data. This resulted in identifying three engagement types; Completing, Disengaging and Sampling. Accordingly, our comparison analysis of the engagement types between the two platforms doesn't consider the Auditing cluster from Coursera in the analysis, as this cluster doesn't appear in our cluster analysis of Edraak data.

Our comparison studies were based on statistical analysis using the chi-square test. Once a significant result was found in an analysis, the adjusted residual helped determine the significant differences that were present. Then, Cramer's V test was used to measure the strength of any identified significant differences between the tested categories.

First, we compared the engagement types proportions of Edraak and Coursera. Our result from the chi-squared test showed that all the engagement types of the three MOOCs on Coursera were significantly different than Edraak's MOOC. However the strength of significance between Edraak's MOOC and Coursera's UG and GS MOOCS was Small, unlike the HS MOOC which showed a Medium strength of significance with Edraak's. Therefore, Edraaks learners could be considered to have similar engagement type proportions to Coursera's UG and GS MOOCs and different proportions to the HS MOOC. This result could be refined by performing more studies that has better control over the sample size.

Then, we looked if there is a difference in the gender proportions between Edraak and Coursera. We found that female enrolment in Edraak was higher that Coursera, whereas male enrolment was less. This result might benefit from a follow-up qualitative research, e.g., survey or interview, to understand the motivations and the objectives of the learners' enrolment.

Next, we compared the distribution of learners over countries with different HDI levels. We found that the vast majority of learners from Coursera's three MOOCs were from countries with Very High HDI levels. On the other hand, most of Edraak's learners were from countries with Low and Medium HDI levels. Moreover, Edraak's learners from low HDI countries showed a higher completion rate than Coursera's learners from the same HDI level countries. These results might indicate that learners from less-developed Arabic countries consider MOOCs as a main educational source. It would be interesting to further explore this hypothesis by running more studies.

Finally, we looked at the use of discussion forums between Edraak's and Coursera's learners. Both platforms have a poor use of the discussion forums. However, the small number of participants who used the discussion forums fell into the Completing engagement type. The reasons behind this poor use of discussion forums should be further studied by seeking more information from the learners of the MOOC. Table 6.16 summarises the results of the statistical analyses that were described in this chapter.

In conclusion, our work is one of the few applications of data analytics in Arabic MOOCs and one of the initial studies to compare Arabic and English platforms. Such a work provides insights to not only the Arabic but also global MOOC communities. Comparing Edraak's engagement types with Coursera's has provided an initial picture of high-level engagement patterns of the Arab MOOC learners. However, additional qualitative studies are required to refine our understanding of leaner engagement patterns in Arabic MOOCs.

Table 6.16: Summary of the results statistical analyses

1) Is there a significant difference in the engagement types proportion between Edraak's and Coursera's learners?		
MOOCs compared	Significance with $p < 0.05$	Strength of significance (Cramer's V)
Coursera HS and Edraak JAVA	✔	Medium
Coursera UG and Edraak JAVA	✔	Small
Coursera GS and Edraak JAVA	✔	Small
2.1) Is there a significant difference in the gender proportion between Edraak's and Coursera's learners?		
MOOCs compared	Significance with $p < 0.05$	Strength of significance (Cramer's V)
Coursera HS and Edraak JAVA	✔	Small
Coursera UG and Edraak JAVA	✔	Small
Coursera GS and Edraak JAVA	✔	Small
2.2) Is there a significant difference in the gender proportion between Edraak's and Coursera's Engagement types?		
MOOCs compared	Significance with $p < 0.05$	Strength of significance (Cramer's V)
Coursera HS and Edraak JAVA	✔	Small
Coursera UG and Edraak JAVA	✔	Small
Coursera GS and Edraak JAVA	✔	Small
3.1) Is there a significant difference in the distribution over countries with different HDI levels between Edraak's and Coursera's learners?		
MOOCs compared	Significance with $p < 0.05$	Strength of significance (Cramer's V)
Coursera HS and Edraak JAVA	✔	Medium
Coursera UG and Edraak JAVA	✔	Medium
Coursera GS and Edraak JAVA	✔	Large
3.2) Is there a significant difference in the distribution over countries with different HDI levels between Edraak's and Coursera's Engagement types?		
MOOCs compared	Significance with $p < 0.05$	Strength of significance (Cramer's V)
Coursera HS and Edraak JAVA	N/A	N/A
Coursera UG and Edraak JAVA	N/A	N/A
Coursera GS and Edraak JAVA	✔	Medium*\Large**

N/A = Coursera data not available
* In the "Medium" HDL level countries
** In the "Low" HDL level countries

Chapter 7

Five-dimensional K-means Clustering on The Edraak Data

In Chapter 5, we analysed learner engagement in Edraak using Kizilcec et al.'s method, and we identified three engagement types that representing Edraak's learners: Sampling, Disengaging and Completing. Then, in Chapter 6, we compared our resulting engagement types with those of the Kizilcec et al. study to detect similarities and differences between Edraak's and Coursera's learners. Our comparison showed that Edraak's learners had engagement types similar to those of Coursera's learners from UG and GS MOOCs, and Edraak's learners from low-HDI countries showed a higher completion rate than Coursera's learners.

In this chapter, we will analyse learner engagement with Edraak using Ferguson and Clow's method, as mentioned in our statement on methodology (Chapter 3). We aim to see if the results achieved from Chapter 5 are sensitive to the particular clustering algorithm.

Following Ferguson and Clow's method, we start by identifying an "engagement description" for each learner in Section 7.1. Then , we use these descriptions to cluster Edraak's learners using a multidimensional clustering algorithm to identify the engagement types of Edraak learners in Section 7.2. Finally, Section 7.3 concludes this chapter with a quantitative statistical analysis that compares the resulting engagement types.

7.1 Edraak's Engagement Description: Ferguson and Clow's Method

Ferguson and Clow [2] studied learner engagement with MOOCs in the Futurelearn platform. Following Kizilcec et al.'s method, they formed an engagement description for each learner based on his/her weekly interaction with the MOOC contents. As Futurelearn allow learners to visit new content (e.g., video or text), submit a weekly assessment (on time or late), and participate in the discussion forum by posting comments, Ferguson and Clow used these interactions as variables in their scoring system to compute learner engagement descriptions, see Table 7.1. They argued that when studying a MOOC platform, researchers should create a scoring system that matches the platform features. Therefore, Ferguson and Clow added the "participate in the discussion forum" to their scoring system, which is not included in Kizilcec et al.'s method.

Table 7.1: Ferguson and Clow's scoring system for forming a learner engagement description for the Futurelearn platform

Score	Interpretation
0	No Interaction
1	Visit new content only
2	Post comments only
3	Visit new content and post comments
4	Submit weekly assessment late
5	Visit new content and submit the weekly assessment late
6	Post comments and submit the weekly assessment late
7	Visit new content, post comments and submit the weekly assessment late
8	Submit the weekly assessment on time
9	Visit new content and submit the weekly assessment on time
10	Post comments and submit the weekly assessment on time
11	Visit new content, post comments and submit the weekly assessment on time

At the end of their MOOCs, each learner had an engagement description consisting of a set of numbers, which represented the scores they received every week. These engagement descriptions were then used as an input for a multidimensional K-means clustering algorithm. Ferguson and Clow's study is explained in detail in section 2.3.

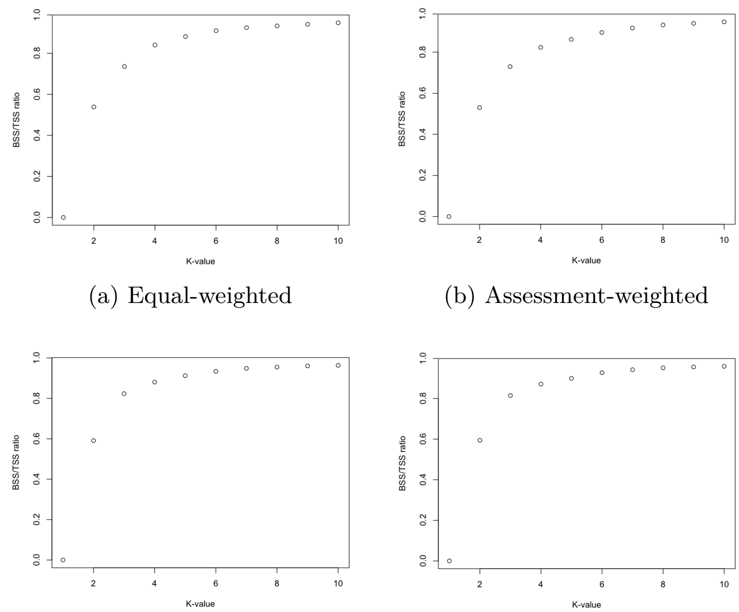
For our analysis of the Edraak data, we emulated Ferguson and Clow's multidimensional clustering approach. However, since our generated data from the Edraak platform contains only two types of interactions, watching video lectures and submitting the weekly assessment, and because late submissions were not allowed, we developed four possible scoring systems to create learners' engagement descriptions. The Equal-weighted scoring system gives a point for each interaction a learner initiates with any MOOC content. The Assessment-weighted scoring system weighs completing the assessment more than watching video lectures. The Diligence-weighted scoring system gives higher scores to learners who watch ALL videos in a week; meaning, for a week with 6 video lectures, watching 1, 2, 3, 4 or 5 videos results in the same score, while watching all 6 is weighted more highly. Finally, the Semi-diligence-weighted scoring system is a way to be more fair to learners, where the more video lectures they watch, the higher score they receive.

We created these scoring systems as an attempt to find the best way to capture the engagement patterns of Edraak's learners, see Table 7.2. For more details on the process of forming these four scoring systems, see Section 3.2.1.

Table 7.2: The interpretations of the four scoring systems used to analyse Edraak's learners

Score	Scoring Systems			
	Equal-weighted	Assessment-weighted	Diligence-weighted	Semi-diligence-weighted
0	No participation	No participation	No participation	No participation
1	Visit videos OR submit assessment	Visit videos only	Visit some videos only	Visit fewer than half of the videos
2	Visit videos AND submit assessment	Submit assessment only	Visit all videos	Visit half or more of the videos but not all
3		Visit content AND submit assessment	Submit assessment only	Visit all videos
4			Visit some videos, submit assessment	Submit assessment only
5			Visit all videos, submit assessment	Visit fewer than half of the videos, submit assessment
6				Visit half or more of the videos but not all, submit assessment
7				Visit all videos, submit assessment

For our multidimensional K-means clustering algorithm (5-dimensional, where 5 represent the number of weeks in our Edraak's MOOC), we explored K-means clustering with the K value varying from 1 to 10 using the R programming language and software environment, and the code can be found in Appendix C. We found that K = 3 is the best number of clusters for each of the four scoring systems. This finding was validated using the NbClust R package, which calculates the optimum number of clusters (K) according to 30 indices [113]. The majority of these indices suggested to cluster Edraak's learners into three clusters, for all scoring systems. In addition, the calculated BSS/TSS ratios of 0.74, 0.73, 0.82 and 0.81 showed a good fit at K= 3, see Figure 7-1. For more information about the validation tests, see Section 3.2.3.



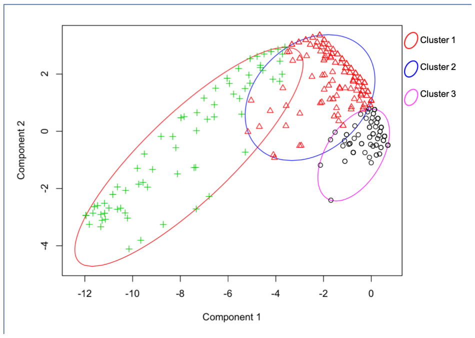
(a) Equal-weighted (b) Assessment-weighted

(c) Diligence-weighted (d) Semi-diligence-weighted

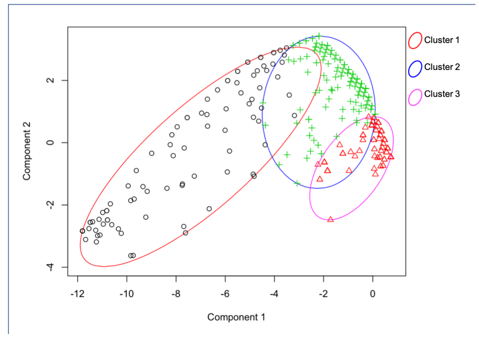
Figure 7-1: BSS/TSS ratio on the Y-axis, and K-means clustering with K={1:10} on the X-axis for the 5-dimensional K-means algorithm

Using "CLUSPLOT", see definition 1 below, we can see that the three clusters obtained from each scoring system were similar, see Figure 8-1. The BSS/TSS ratios suggested that the Diligence-weighted scoring system had the best fit (0.82). Figure 8-1 also displays the overlap between the clusters obtained from each scoring system, and we can see that the Diligence-weighted scoring system had the least overlap. Based on this result, we chose the three clusters obtained from the Diligence-weighted scoring system to represent Edraak's learners. Appendix E contains the detailed engagement types of the clusters from each scoring system.

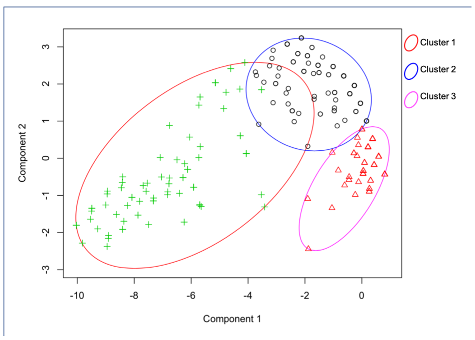
Definition 1. *CLUSPLOT is a graphical display that plots multidimensional clusters in two dimensional using principal component analysis (PCA) [121]. PCA transfers a dataset with multiple dimension and plot it into two-dimensions, while keeping most of the variation present in the dataset [122]. It creates two new dimensions, component 1 and component 2 in figure 8-1, to represent the differences between the dataset's multiple dimensions.*



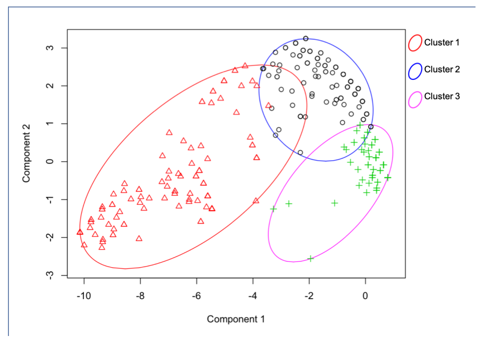
(a) Equal-weighted



(b) Assessment-weighted



(c) Diligence-weighted



(d) Semi-diligence-weighted

Figure 7-2: Clusplot of Edraak data using the four scoring systems. Similar clusters (ellipses) are highlighted in the same colour

Remark. *Four scoring systems were used to compute learner's engagement descriptions for the Edraak MOOC. All systems used two types of interactions with the MOOC contents: watching video lectures and submitting a weekly assessment. Within all four scoring systems, clustering the Edraak data with K=3 showed a good fit. The Diligence-weighted scoring system was used for clustering the Edraak data, as it showed the minimum overlap between the three clusters.*

7.2 Edraak's Engagement Types: Ferguson and Clow's Method

Our 5-dimensional K-means clustering algorithm classified Edraak's learners into three clusters that represent their engagement types. We analysed the learner interaction with the MOOC contents to define each engagement type.

Cluster 1: The largest cluster, with 2202 learners. Fewer than half of them watched at least one video lecture at the beginning of the MOOC before joining the majority who dropped out by the following week. In addition, learners from this cluster had no engagement with the weekly assessments throughout the MOOC, see Figure 7-3. The engagement in this cluster is similar to the "Samplers" engagement from Ferguson and Clow [2], therefore, we called this engagement type **Samplers**. For more on the detailed engagement of Cluster 1, see Tables 21 and 22 in Appendix E.3.

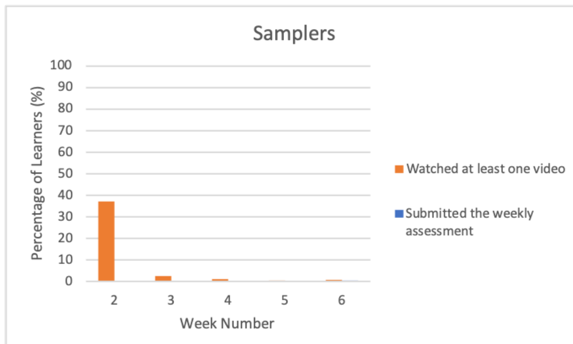


Figure 7-3: Learner interaction with the Edraak MOOC contents for the Samplers engagement type

Cluster 2: Consisted of 350 out of 2736 learners. They showed great engagement with the contents at the beginning of the MOOC before the quickly decreased, see Figure 7-4. The majority of learners dropped out halfway through the MOOC, which is similar to the "Mid-way Dropouts" engagement type found in Ferguson and Clow [2], therefore, we called this engagement type **Mid-way Dropouts**. For more on the detailed engagement of Cluster 2, see Tables 23 and 24 in Appendix E.3.

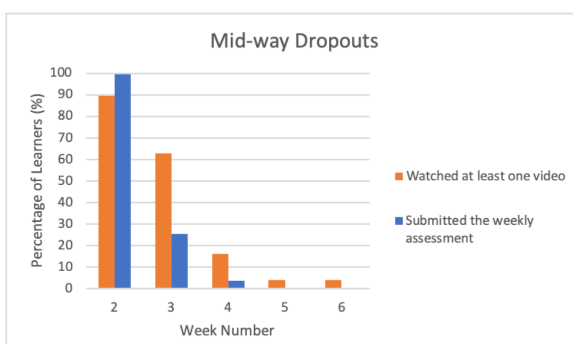


Figure 7-4: Learner interaction with the Edraak MOOC contents for the Mid-way Dropouts engagement type

Cluster 3: This cluster had 184 learners, who maintained good interaction with the weekly assessments at the first half of the MOOC before this interaction slightly decreased towards the end of the MOOC. However, learners from this cluster did not have a similar interest towards video lectures, as shown in Figure 7-5. We called this engagement type **Completers**, as they were the only cluster to complete the MOOC. For more on the detailed engagement of Cluster 3, see Tables 25 and 26 in Appendix E.3.

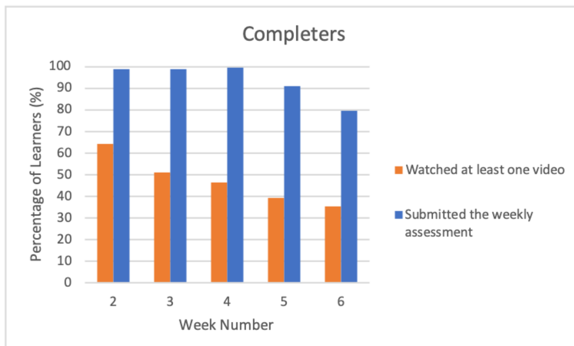


Figure 7-5: Learner interaction with the Edraak MOOC contents for the Completers engagement type

7.3 Quantitative Analysis

Three engagement types were obtained from our 5-dimensional K-means clustering algorithm: Samplers, Mid-way Dropouts and Completers. We compared the demographics of these three engagement types, as we did with the engagement types obtained from the 1-dimensional algorithm in Section 5.3. Comparing the three engagement types from our 5-dimensional algorithm did not output any different result than what we obtained for the three engagement types from our 1-dimensional algorithm. The full analysis can be found in Appendix F.

This result means that both clustering algorithms resulted in the same three engagement types. To prove that finding, we compared the following engagement types, from the two clustering algorithms, as they share the same descriptions: Sampling and Samplers, see Table 7.3, Disengaging and Mid-way Dropouts, see Table 7.4, and Completing and Completers, see Table 7.5. A chi-squared test showed **no significant** difference between the compared groups, see Appendix G.

Table 7.3: The demographics of Edraak's Sampling and Samplers

		Cluster 1	
		Sampling	Samplers
Number of learners		2302	2202
Gender	Males	67.15%	67.35%
	Females	32.71%	32.52%
Qualification	No formal education	0.83%	0.86%
	Junior high school	4.91%	5%
	High school	31.63%	31.74%
	Associate	3.52%	3.59%
	BSc	43.4%	43.1%
	MSc	12.03%	11.94%
	PhD	0.56%	0.54%
	Other	3.13%	3.22%
HDI level	Very high	14.20%	14.12%
	High	23.58%	23.48%
	Medium	51.47%	51.54%
	Low	9.47%	9.63%
Mean number of comments		0.003	0.003
Mean number of videos watched		1.19	1.10

Table 7.4: The demographics of Edraak's Disengaging and Mid-way Dropouts

		Cluster 2	
		Disengaging	Mid-way Dropouts
Number of learners		254	350
Gender	Males	69.68%	67.43%
	Females	30.31%	32.57%
Qualification	No formal education	0%	0%
	Junior high school	6.7%	6%
	High school	35.83%	34.29%
	Associate	4.33%	3.71%
	BSc	41.73%	44.29%
	MSc	7.87%	9.14%
	PhD	0%	0.29%
	Other	3.54%	2.29%
HDI level	Very high	14.17%	14.57%
	High	29.13%	27.71%
	Medium	48.81%	50%
	Low	6.69%	6.57%
Mean number of comments		0.01	0.01
Mean number of videos watched		12.02	9.37

Table 7.5: The demographics of Edraak's Completing and Completers

		Cluster 3	
		Completing	Completers
Number of learners		180	184
Gender	Males	76.66%	77.17%
	Females	23.33%	22.83%
Qualification	No formal education	0%	0%
	Junior high school	3.34%	2.72%
	High school	34.44%	33.7%
	Associate	2.78%	2.72%
	BSc	43.89%	43.48%
	MSc	8.89%	9.78%
	PhD	2.22%	2.17%
	Other	4.44%	5.43%
HDI level	Very high	18.33%	18.48%
	High	25%	26.09%
	Medium	43.33%	41.85%
	Low	12.22%	11.96%
Mean number of comments		0.16	0.16
Mean number of videos watched		14.58	14.79

7.4 Discussion

Defining the Edraak learner features would enable MOOC designers to develop better platforms that meet learners' needs. In Chapter 5, we analysed Edraak's learners using a 1-dimensional K-means clustering algorithm that was used by Kizilcec et al. [1] to analyse MOOCs from the Coursera platform. This algorithm showed that Edraak's learners fall into three engagement types: Sampling, Disengaging and Completing. In 2015, Ferguson and Clow [2] used the same approach as that of Kizilcec et al. [1] to analyse MOOCs from the Futurelearn platform. However, they could not identify clear engagement types, and they concluded that the 1-dimensional approach discarded useful information. Therefore, they modified Kizilcec et al.'s algorithm by converting it to use multidimensional K-means clustering, which helped them to identify the engagement types of Futurelearn's learners. In this part of the study, we considered whether incorporating this modification to Kizilcec et al.'s approach would give us a different, or even clearer, definition of the engagement types of Edraak's learners.

Kizilcec et al.'s clustering approach started by assigning weekly scores to learners based on completing a weekly assessment first, which grants the highest score, and then based on the interaction with video lectures if the learners failed to complete the assessment. Ferguson and Clow [2] suggested that differences in the platform design should be considered when assigning weekly scores, meaning that a suitable scoring system should be used to analyse MOOCs from different platforms. As the Futurelearn platform incorporates discussion in their learning approach, learners weekly participation in the discussion forum was included in the scoring system. We considered Ferguson and Clow's suggestion when designing our scoring system; however, Edraak's platform provided us with only limited data. The only variables that we could use in our scoring system were watching the video lectures and completing the assessment.

Using Ferguson and Clow's approach with only two variables (videos and assessment), we designed four different scoring systems with the aim to find the one best suited for Edraak's platform. The Equal-weighted scoring system weighs the two variables equally, whereas the Assessment-weighted scoring system weighs completing the assessment by one point more. The Diligence-weighted scoring system rewards learners who watch all the video lectures with the highest score, and the Semi-diligence-weighted scoring system distinguishess learners based on how many of the overall video lectures they watch. Section 3.2.1 explained the four scoring systems in detail with examples. Our four scoring systems cover all possible interactions with the MOOC contents, unlike our first analysis that followed Kizilcec et al.'s scoring system, where we ignored the interaction with the video lectures once a learner completed the weekly assessment. Analysing the Edraak data using both approaches, those of Kizilcec et al. [1] and Ferguson and Clow [2], resulted in the same engagement types using the two variables (videos and assessment). We suggest that this finding occurred due to relying on only two variables to understand learner engagement with Edraak's MOOCs. This approach is more likely to produce differing results when including more variables, as seen with Ferguson and Clow's [2] work.

In conclusion, although analysing the Edraak data using Ferguson and Clow's [2] approach resulted in the same three engagement types that we obtained using Kizilcec et al.'s approach, we conclude that Ferguson and Clow's approach has advantages when considering the platform design in their analysis and clustering their data with a multidimensional K-means clustering algorithm. For future analysis, to understand learner engagement in any MOOC platform, the use of Ferguson and Clow's approach would benefit from using more than two variables.

Chapter 8

Comparison Between Edraak's and Futurelearn's learners

Previously, we analysed Edraak's learners engagement with MOOCs, using Kizilcec et al. [1] method that was used to analyse learners engagement in the Coursera platform, see Chapter 5. Then, we compared our resulted engagement types with theirs, in Chapter 6. The aim was to understand how Arabic-speaking learners engage with MOOCs? and how it is similar to or different from the engagement of English-speaking learners?

In Chapter 7, we made some changes in our analysis of Edraak's learners, by followed Ferguson and Clow [2] method that was used to analyse learners engagement in the Futurelearn platform. However, we had the same results that we obtained from our first analyses, and we concluded that we need more data from Edraak platform that describes learners weekly activities in order to understand their engagement types. In an attempt to further understand how Arabic-speaking learners engagement with MOOCs is similar to or different from the engagement of English-speaking learners? this Chapter compares Edraak's and Futurelearn's engagement types.

Ferguson and Clow [2] analysed learners engagement with four MOOCs from the Futurelearn platform, using a multidimensional K-means clustering algorithm. They identified seven engagement types: Samplers, Strong Starters, Returners, Mid-way Dropouts, Nearly There, Late Completers and Keen Completers, see Section 2.3.3. Due to differences in the number of weeks in each MOOC and the number of assessment, the Returners engagement type existed in 3 out of the 4 analysed MOOCs, and the Mid-way Dropouts engagement type existed in 2 out of the 4 analysed MOOCs. For more information on Ferguson and Clow [2] work, see section 2.3.

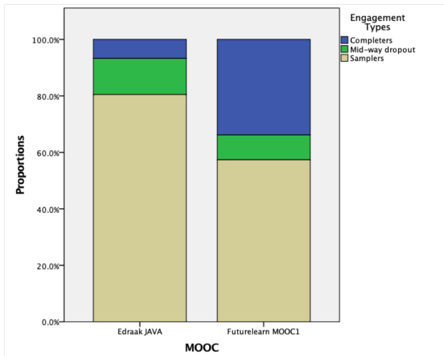
When we analysed learners engagement with a MOOC from the Edraak platform, using a multidimensional K-means clustering algorithm, we identified three engagement types: Samplers, Mid-way Dropouts and Completers, see Section 7.2. We compared our resulted engagement types with the engagement types of each of the four MOOCs from Ferguson and Clow [2] work. The aim is to identify similarities or differences in the engagement with MOOCs between Arabic- and English-speaking learners, and to see if the result are similar to our comparison with Coursera in Chapter 6.

The **chi-square** test was used to compare the engagement types between the two platforms, Edraak and Futurelearn, considering the significance levels shown in Table 8.1. For more information about the chi-squared test see Sections 3.3.1.

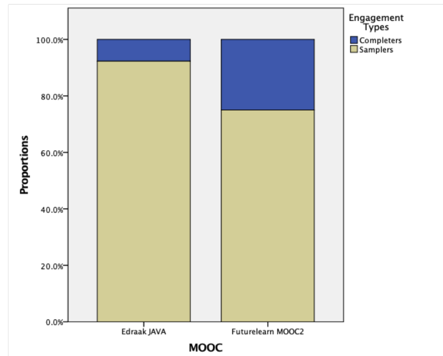
Table 8.1: The considered significance levels for the used statistical tests

Statistical Tests	Significance Level
Chi-Square	$p < 0.05$
Adjusted Residuals	± 1.96
Cramer's V	See table 3.6 in section 3.3.1.2

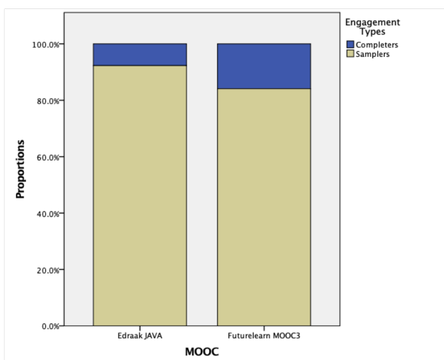
8.1 Proportions of Engagement Types



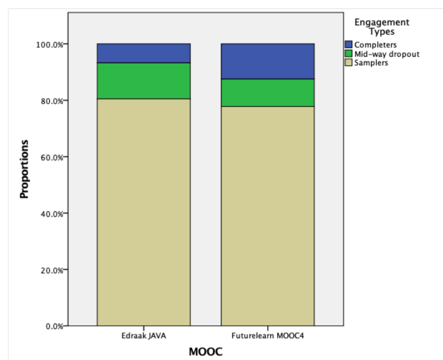
(a) Futurelearn's Physical Sciences MOOC and Edraak's JAVA MOOC



(b) Futurelearn's Life Sciences MOOC and Edraak's JAVA MOOC



(c) Futurelearn's Arts MOOC and Edraak's JAVA MOOC



(d) Futurelearn's Business MOOC and Edraak's JAVA MOOC

Figure 8-1: The proportions of the engagement types of learners in Futurelearn's four MOOCs in comparison to Edraak's JAVA MOOC

Edraak and Futuerlearn have three common engagement types Samplers, Mid-way Dropouts and Completers (or Keen Completers for Futuerlearn). We compared the engagement types of Edraak's MOOC (JAVA) with the engagement types of Futurelearn's MOOC (Physical Sciences). Figure 8-1a shows the engagement types proportions between the two engagement types. We can notice the large proportion of Completers in the futurelearn platform on the contrast to Edraak, which had higher proportion of non-completers (Samplers and Mid-way Dropouts). Chi-squared test shows that the proportion of Completers in Futurelearn is **significantly higher** than Edraak. In addition, the proportions of Samplers and Mid-way Dropouts in Futurelearn is **sig-**

nificantly lower than Edraak (X^2 = 1784.199, *df* = 6, $p < .001$), see Table 8.2. Cramer's V test shows that the strength of significance is **Large** with value of 0.480, which means that this significance is considered acceptable [116].

Running a similar comparison with the second MOOC from Futurelearn, Life Sciences MOOC; see Figure 8-1b, showed the proportions of the only two engagement types that existed in both Edraak and Futurelearn platforms. Similar to the comparison with the first Futurelearn MOOC, Physical Sciences MOOC, chi-squared test shows that Futurelearn platform has a **significantly higher** proportion of Completers and a **significantly lower** proportion of Samplers compared with Edraak platform (X^2 = 2224.833, *df* = 7, $p < .001$), see Table 8.3. Cramer's V test showed a **Large** strength of significance with value of 0.610, which makes this significance is considered acceptable [116]. A very similar result was found when comparing Edraak MOOC with the third MOOC from Futurelearn, Arts MOOC, see Figure 8-1c. Chi-squared value was X^2 = 3409.998, *df* = 7, $p < .001$ and a Cramer's V value of 0.467, see Table 8.4.

We Also compared Edraak MOOC engagement types proportions with the ones of the Business MOOC on Futurelearn, see Figure 8-1d. Chi-squared test showed that Futurelearn platform has a **significantly higher** proportion of Completers and a **significantly lower** proportion of Mid-way dropouts and Samplers compared with Edraak platform (X^2 = 1074.875, *df* = 6, $p < .001$), see Table 8.5. This significance is considered acceptable as Cramer's V test showed a **Large** strength of significance with value of 0.293 [116].

Table 8.2: Engagement Types * MOOCs Crosstabulation: Futurelearn's Physical Sciences MOOC and Edraak's JAVA MOOC

		Physical Sciences	JAVA	Total
Samplers	Count	1977	2202	4179
	Expected Count	2704.6	1474.4	4179
	% within platform	39.4%	80.5%	53.9%
	Adjusted Residual	-34.7	34.7	
Mid-way dropouts	Count	304	350	654
	Expected Count	423.3	230.7	654
	% within platform	6.1%	12.8%	8.4%
	Adjusted Residual	-10.2	10.2	
Completers	Count	1166	184	1350
	Expected Count	873.7	476.3	1350
	% within platform	23.2%	6.7%	17.4%
	Adjusted Residual	18.3	-18.3	
Strong starters	Count	558	0	558
	Expected Count	361.1	196.9	558
	% within platform	11.1%	0%	7.2%
	Adjusted Residual	18.1	-18.1	
Returners	Count	304	0	304
	Expected Count	196.7	107.3	304
	% within platform	6.1%	0%	3.9%
	Adjusted Residual	13.1	-13.1	
Nearly there	Count	304	0	304
	Expected Count	196.7	107.3	304
	% within platform	6.1%	0%	3.9%
	Adjusted Residual	13.1	-13.1	
Late completers	Count	406	0	406
	Expected Count	262.8	143.2	406
	% within platform	8.1%	0%	5.2%
	Adjusted Residual	15.3	-15.3	
Total	Count	5019	2736	7755
	% within platform	100%	100%	100%

Table 8.3: Engagement Types * MOOCs Crosstabulation: Futurelearn's Life Sciences MOOC and Edraak's JAVA MOOC

		Life Sciences	JAVA	Total
Samplers	Count	1263	2202	3465
	Expected Count	1878.1	1586.9	3465
	% within platform	39%	80.5%	58%
	Adjusted Residual	-32.4	32.4	
Mid-way dropouts	Count	0	350	350
	Expected Count	189.7	160.3	350
	% within platform	0%	12.8%	5.9%
	Adjusted Residual	-21	21	
Completers	Count	421	184	605
	Expected Count	327.9	277.1	605
	% within platform	13%	6.7%	10.1%
	Adjusted Residual	8	-8vv	
Strong starters	Count	453	0	453
	Expected Count	245.5	207.5	453
	% within platform	14%	0%	7.6%
	Adjusted Residual	20.4	-20.4	
Returners	Count	259	0	259
	Expected Count	140.4	118.6	259
	% within platform	8%	0%	4.3%
	Adjusted Residual	15.1	-15.1	
Nearly there	Count	194	0	194
	Expected Count	105.2	88.8	194
	% within platform	6%	0%	3.2%
	Adjusted Residual	13	-13	
Late completers	Count	227	0	227
	Expected Count	123	104	227
	% within platform	7%	0%	3.8%
	Adjusted Residual	14.1	-14.1	
Supplementary cluster	Count	421	0	421
	Expected Count	228.2	192.8	421
	% within platform	13.0%	0%	7%
	Adjusted Residual	19.6	-19.6	
Total	Count	3238	2736	5974
	% within platform	100%	100%	100%

Table 8.4: Engagement Types * MOOCs Crosstabulation: Futurelearn's Arts MOOC and Edraak's JAVA MOOC

		Arts	JAVA	Total
Samplers	Count	5964	2202	8166
	Expected Count	6739.5	1426.5	8166
	% within platform	46.1%	80.5%	52.1%
	Adjusted Residual	-32.7	32.7	
Mid-way dropouts	Count	0	350	350
	Expected Count	288.9	61.1	350
	% within platform	0%	12.8%	2.2%
	Adjusted Residual	-41.1	41.1	
Completers	Count	1128	184	1312
	Expected Count	1082.8	229.2	1312
	% within platform	8.7%	6.7%	8.4%
	Adjusted Residual	3.4	-3.4	
Strong starters	Count	1289	0	1289
	Expected Count	1063.8	225.2	1289
	% within platform	10%	0%	8.2%
	Adjusted Residual	17.2	-17.2	
Nearly there	Count	967	0	967
	Expected Count	798.1	168.9	967
	% within platform	7.5%	0%	6.2%
	Adjusted Residual	14.8	-14.8	
Late completers	Count	32	0	32
	Expected Count	26.4	5.6	32
	% within platform	0.2%	0%	0.2%
	Adjusted Residual	2.6	-2.6	
Supplementary cluster	Count	3224	0	3224
	Expected Count	2660.8	563.2	3224
	% within platform	24.9%	0%	20.6%
	Adjusted Residual	29.3	-29.3	
Supplementary cluster 2	Count	322	0	322
	Expected Count	265.7	56.3	322
	% within platform	2.5%	0%	2.1%
	Adjusted Residual	8.3	-8.3	
Total	Count	12926	2736	15662
	% within platform	100%	100%	100%

Table 8.5: Engagement Types * MOOCs Crosstabulation: Futurelearn's Business MOOC and Edraak's JAVA MOOC

		Business	JAVA	Total
Samplers	Count	5476	2202	7678
	Expected Count	5999.3	1678.7	7678
	% within platform	56%	80.5%	61.4%
	Adjusted Residual	-23.2	23.2	
Mid-way dropouts	Count	684	350	1034
	Expected Count	807.9	226.1	1034
	% within platform	7%	12.8%	8.3%
	Adjusted Residual	-9.7	9.7	
Completers	Count	880	184	1064
	Expected Count	831.4	232.6	1064
	% within platform	9%	6.7%	8.5%
	Adjusted Residual	3.8	-3.8	
Strong starters	Count	978	0	978
	Expected Count	764.2	213.8	978
	% within platform	10%	0%	7.8%
	Adjusted Residual	17.2	-17.2	
Returners	Count	684	0	684
	Expected Count	534.5	149.5	684
	% within platform	7%	0%	5.5%
	Adjusted Residual	14.2	-14.2	
Nearly there	Count	489	0	489
	Expected Count	382.1	106.9	489
	% within platform	5%	0%	3.9%
	Adjusted Residual	11.9	-11.9	
Late completers	Count	587	0	587
	Expected Count	458.7	128.3	587
	% within platform	6%	0%	4.7%
	Adjusted Residual	13.1	-13.1	
Total	Count	9778	2736	12514
	% within platform	100%	100%	100%

8.2 Discussion

As an extension to our work in Chapter 6, where we compared Coursera's and Edraak's learners to identify similarities and differences in the engagement patterns of Arabic- and English-speaking learners with MOOCs, in this chapter, we compared the engagement patterns of Edraak's and Futurelear's learners. This might provide a further understanding of Edraak learners, and accordingly improve the platform design.

Ferguson and Clow [2] studied learners engagement with four MOOCs from the Futurelearn platform and they identify seven engagement types that represent Futurelearn's learners: Samplers, Strong starters, Returners, Mid-way dropouts, Nearly there, Late completers and Keen completers, see Section 2.3.3. Due to differences in the number of assessments and in the duration of the four MOOCs, "Returners" engagement type existed in 3 out of the 4 MOOCs, whereas "Mid-way dropouts" engagement type existed in 2 out of the 4 MOOCs. To compare the engagement patterns of Edraak with Futurelearn, we applied their clustering approach on Edraak's data. This resulted in identifying three engagement types: Samplers, Mid-way dropouts and Completers. Accordingly, our comparison analysis of the engagement types between the two platform focuses only on comparing the engagement types that exist in the compared MOOCs.

The comparison of the engagement types between the two platforms, Edraak and Futurelearn, was performed using chi-square test. As explained in Section 3.3.1, if the test shows a significant result, then the value of the adjusted residual identifies which groups are significantly different. Then, Cramer's V test determine the strength of this significance. We found that Futurelearn learners has a higher proportion of Completers, and a lower proportion of Samplers compared with Edraak learners. This is the only difference that we could identified, as the data obtained from Ferguson and Clow's study was limited.

Looking back at the result of comparing the engagement types of Edraak's learners with Coursera's learners in Section 6.1, and linking it with the comparison between Edraak's and Futurelearn's engagement types, we observed that Edraak's and Coursera's learners engage with MOOCs in a similar way. Both platforms have a high dropout rate and low completion rate. On the contrast, the completion rate in Futurelearn seems to be higher than Coursera and Edraak.

Ferguson and Clow [2] attempted to analyse Futurelearn learners by coping Kizilcec et al.'s [1] approach. However, they failed to identify clear engagement types of their learners. On their second attempt, they benefited from using a multidimensional K-means clustering algorithm. In addition, they benefited from considering the platform pedagogy into their analysis, as Futurelearn platform employs the "social-constructivist pedagogy" as an educational approach, which encourages discussion among learners themselves and with the educators, see Section 2.3.2.1. Since we also used a multidimensional K-means clustering algorithm to analyse Edraak data that showed the same result we had from our analysis using a 1-dimensional K-means clustering algorithm, we believe that incorporating the social-constructivist pedagogy in the Futurelearn platform design might had an association with obtaining a higher completing rate than Coursera and Edraak.

Chapter 9

Conclusion

In this research, we investigated the adoption of massive open online courses (MOOCs) in Arabic-speaking countries. The aim was to identify the characteristics of Arabic-language MOOC learners and to compare them with learners from English-language MOOCs. The objective was to undertake an analysis of learner actions to help inform MOOC providers and designers in the development of their learning platforms. It would also help them to develop and improve MOOC platforms accordingly. Moreover, this research will help educators and researchers to plan and conduct further analyses, as the field of MOOC engagement analysis lacks such research that targets Arabic MOOC platforms and Arabic-speaking learners.

To achieve this aim, we requested learners' data from an Arabic-language MOOC platform called Edraak. We analysed these data using the two clustering approaches presented by Kizilcec et al. [1] and Ferguson and Clow [2], who analysed learners' data from the Coursera and Futurelearn MOOC platforms, respectively. Then, we compared our resulted clusters with the ones from Kizilcec et al.'s and Ferguson and Clow's studies. The approaches were different in two respects:

1. **Scoring:** In the first approach we adopted a simple scoring system presented by Kizilcec et al. [1], which was based on completing the weekly assessment first, then on watching video lectures. In the second approach, we followed Ferguson and Clow's [2] suggestion by creating a scoring system that suits the Edraak platform's design, and we considered all possible interactions with the MOOC

contents.

2. **Clustering:** Following Kizilcec et al. [1], our clustering in the first approach was one-dimensional, where we used the L_1 norm equation to sum the scores on each engagement description into a single digit. However, for the second approach we followed Ferguson and Clow [2] and skipped the use of the L_1 norm equation, making our clustering five-dimensional.

The two approaches used in this research are detailed in section 3.2. The following sections summarise the research findings and present a general discussion regarding the adoption of MOOC in the Arabic-speaking countries in comparison to English-speaking countries. Finally, this chapter will conclude with a section listing the research limitations and suggested future work.

9.1 Summary of Research Findings

Before investigating the Arabic learners' engagement with MOOCs, we started our analysis identifying the Arabic-speaking MOOC learners' profiles. Using the data gathered and provided to us by the Edraak MOOC platform, we found the following:

1. Males have a higher enrolment rate than females.
2. Bachelor's degree holders followed by high-school degree holders represent the majority of learners.
3. The majority of male enrolments came from countries with Medium HDI level, whereas the majority of female enrolments came from Very High HDI level countries.
4. Poor use of the discussion forum in general, despite the higher mean number of comments from males compared to females, was prevalent.

When studying the Arabic learners' engagement with MOOCs, applying the approaches of both Kizilcec et al. [1] and Ferguson and Clow [2] resulted in clustering Edraak's learners into three clusters that represent three types of engagement (see Table 9.1).

Table 9.1: The three resulting clusters of Edraak's learners after applying the approaches of Kizilcec et al. and Ferguson and Clow

	Approach Followed	Cluster Name	Description
Cluster 1	Kizilcec et al.	Sampling	Consisted of learners who had no interaction with the MOOC contents, or dropped out from the MOOC after the first week.
	Ferguson and Clow	Samplers	
Cluster 2	Kizilcec et al.	Disengaging	Consisted of learners who had a good interaction with the MOOC contents at the first week of the MOOC, and gradually dropped out by the middle of the course.
	Ferguson and Clow	Mid-way Dropouts	
Cluster 3	Kizilcec et al.	Completing	Consisted of learners who had a good interaction with the MOOC contents, especially the weekly assessments, throughout the MOOC duration.
	Ferguson and Clow	Completers	

Comparing the three engagement types of Edraak, we identified differences in the interactions with the discussion forum and the video lectures. Learners who fall into the Completing/Completers group have a higher engagement with the discussion forum than learners from the other groups. As for watching the posted video lectures, we found the Sampling/Samplers group to have the lowest interaction. On the similarity side, Edraak's three engagement types had similar proportions of gender, qualification and HDI levels of the learners countries. Then, we compared Edraak's learners with Coursera's learners from Kizilcec et al.'s work [1], and we identified the following similarities and differences:

Similarities	1. Edraak's engagement types proportions are similar to Coursera's engagement types for the undergraduate and graduate level MOOCs.
	2. Edraak's learners have a similar gender proportion to Coursera's learners in both the overall enrolment and within each engagement type.
	3. Edraak's and Coursera's learners have a similarly low usage of the discussion forum.
Differences	1. Edraak's engagement types proportions are different from Coursera's engagement types for the high-school level MOOC.
	2. Edraak's learners are mostly from countries with a medium HDI level, whereas the majority of Coursera's learners are from countries with a very high HDI level.
	3. Edraak has more "Completing" learners from countries with high, medium and low HDI levels than Coursera, where the majority of the "Completing" are from countries with a very high HDI level.

Then we compared the engagement types of Edraak's learners with the ones of Futurelearn's learners from Ferguson and Clow's work [2]. This comparison had essentially one result, stating that Futurelearn MOOCs had a significantly higher proportion of Completer learners and a significantly lower proportion of Sampler learners than Edraak MOOC.

9.2 Limitation

This research relies on obtaining data from Arabic MOOC platforms and, as we mentioned before in Section 1.2, Arabic MOOC platforms started launching in 2013, and they are gradually gaining in reputation among Arabic learners. However, the platforms' founders and their teams are more interested in designing courses and launching new platforms than in collecting usage data from their learners or analysing platforms' activities. We wanted to obtain a sample that represents learners from around the Arabic world; therefore our focus was on the two leading Arabic MOOC platforms, Edraak and Rwaq.

The communication with the Edraak team was not easy at firs, as there was limited contact information available on the platform. Although the team was keen to help us and supportive to our research, their platform was fairly new and they were working on improving it. This led to a limited amount of data gathered and slowed the analysis process, as the Edraak team was waiting to complete recruitment on their side to manage their data. Then we contacted Rwaq to request data from their platform; however, the same problem with data management existed.

9.3 General Discussion

In this research, we wanted to identify learners' engagement types in the Arabic-speaking MOOC, and than undertake a comparison with the English-speaking MOOCs from leading platforms. For this purpose we analysed two papers that followed similar approaches to study learners on Coursera and Futurelearn platforms. When studying learners' engagement with Coursera's MOOCs, Kizilcec et al. [1] created a simple classification method based on two elements of interactions: video lectures and assessment.

They believe the simplicity of their method will allow it to be applicable to other platforms, regardless of the platform's pedagogical approach or the contents they provide. We found this to be valid when we used their method to analyse Edraak's data. The only variables that we got from Edraak's team were video lectures and assessments. With limited variables, simplifying the classification method resulted in well fitted clusters, as seen in Sections 5.1 and 5.2. By contrast, Ferguson and Clow [2] found that this simple classification method does not suite MOOCs on the Futurelearn platform. Therefore, they created a more complex classification method that included the platform's pedagogical approach and additional interactions options. Their result provided well described engagement patterns for Futurelearn's learners; however, the downside of a complex method is that it is inapplicable in different MOOC platforms. Testing whether a complex classification method could be created that is only applicable on Edraak's MOOC was not possible due to the limitation in data generation from Edraak platform, see Section 9.2.

Comparing Arabic-speaking MOOC learners with English-speaking MOOC learners was not a simple task, as English-speaking learners had different engagement types depending on the MOOC platform used. (See Sections 2.2.3 and 2.3.3 for the engagement types of Coursera's and Futurelearn's MOOCs, respectively). Our comparison showed that the engagement of Edraak's learners with MOOCs are closer to Coursera's learners than to Futurelearn's learners. The limited variables and the use of a simple classification method might have led to this result. However, Ferguson and Clow [2] related the difference between Coursera's and Futurelearn's learners to the difference in the pedagogical approaches adopted by the two platforms. (See Section 1.3 for the platforms' pedagogies). Futurelearn's social-constructivist pedagogy encourages the learners to communicate, share and discuss ideas to gain knowledge rather than relying only on the MOOC teacher, which helped learners to stay connected throughout the course. Kizilcec et al. [1], on the other hand, reported a low participation rate in the discussion forums when analysing Coursera's MOOCs, which follow the cognitive-behaviourist pedagogy. This pedagogy uses the MOOC teacher as the main source of knowledge and the discussion between learners is usually focused on the received knowledge from the MOOC.

There is no clear pedagogical approach that is adopted by the Arabic MOOC platform. The interest is mainly toward launching new platforms and providing MOOCs that cover most subjects (as noted by the founders of Rwaq platform when they launched two more platforms that target specific learners, as discussed in Section 1.2). Arabic MOOC platforms provide discussion forums; however, they are rarely used for sharing or exchanging knowledge with the educational community. Looking back at the three types of pedagogical approaches, section 1.3, we cannot identify which one of them describes the pedagogical approach of the Arabic MOOC platforms. At this stage, Arabic MOOC platforms are designed as traditional classrooms in an online platform where the teacher provides the knowledge and learners are expected to study it and apply it to the weekly assessment to pass the MOOC. Therefore, it might be beneficial to investigate links between learners behaviour in traditional classrooms and MOOC courses to better understand the resulted engagement types of Arabic learners. In summary, MOOC platforms can be a great learning tool that benefits many learners without the restrictions of time, place and usually money; however, this will be difficult to achieve without having a clear pedagogical approach to follow. We encourage MOOC providers to give more attention to educational pedagogies when designing MOOCs, as done by the Futurelearn platform. In addition, enhancing data management in Arabic platforms is important for tracking user/learner activity and gathering data for future research.

9.4 Future Work

This research provided a statistical analyses of engagement in Arabic language MOOCs. The results and conclusion of this research represent an initial step in understanding and accordingly improving Arabic learners' MOOC experience, as well as providing guidance for further analyses. For future research, obtaining more data from the Arabic MOOC platforms will enable the creation of a more complex classification method. Examples of such data are age, employment status, weekly comments, assessment submission time, video lecture watching time, enrolment time and requesting a completion certificate. This research can also help future studies that are interested in gathering data from learners by directly using qualitative data/survey. Looking back to our results, we have pointed toward some interesting questions, such as the following:

- What caused the quick dropout of the "Sampling" learners?
- What was the objective of the "Sampling" learners' enrolment?

- Why did the “Disengaging” learners return to the course after the first week, unlike the “Sampling” learners?
- What was the objective of the “Disengaging” learners’ enrolment?
- What was the motivation behind completing the course for the “Completing” learners?
- Were the “Completing” learners interested in requesting a completion certificate? If so, why?
- What are the reasons behind the poor use of discussion forums?

Other research directions might include investigating the effect of MOOC’s pedagogical approaches on Arabic learners’ engagement. This can be done by collaborating with an Arabic MOOC platform to design a MOOC that adopts a specific pedagogical approach, and measuring the changes in the learners’ engagement.

Chapter 10

References

[1] R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses," in *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, ser. LAK '13. New York, NY, USA: ACM, 2013, pp. 170–179. [Online]. Available: http://doi.acm.org/10.1145/2460296.2460330

[2] R. Ferguson and D. Clow, "Examining engagement: Analysing learner subpopulations in massive open online courses (moocs)," in *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, ser. LAK '15. New York, NY, USA: ACM, 2015, pp. 51–58. [Online]. Available: http://doi.acm.org/10.1145/2723576.2723606

[3] A. M. F. Yousef, M. A. Chatti, U. Schroeder, M. Wosnitza, and H. Jakobs, "A review of the state-of-the-art," *Proceedings of CSEDU*, pp. 9–20, 2014.

[4] S. Mulik, N. Yajnik, and M. Godse, "Determinants of acceptance of massive open online courses," in *2016 IEEE Eighth International Conference on Technology for Education (T4E)*. IEEE, 2016, pp. 124–127.

[5] B. Wu and X. Chen, "Continuance intention to use moocs: Integrating the technology acceptance model (tam) and task technology fit (ttf) model," *Computers in Human Behavior*, vol. 67, pp. 221–232, 2017.

[6] Y. Ouyang, C. Tang, W. Rong, L. Zhang, C. Yin, and Z. Xiong, "Task-technology fit aware expectation-confirmation model towards understanding of moocs con-

tinued usage intention," *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.

[7] A. McAuley, B. Stewart, G. Siemens, and D. Cormier, "The mooc model for digital practice," 2010.

[8] F. de Langen and H. van den Bosch, "Massive open online courses: Disruptive innovations or disturbing inventions?" *Open Learning: The Journal of Open, Distance and e-Learning*, vol. 28, no. 3, pp. 216–226, 2013.

[9] K. R. Clark, B. L. Vealé, and L. K. Watts, "A review of the use of massive open online courses (moocs) in medical imaging education," *Internet Journal of Allied Health Sciences and Practice*, vol. 15, no. 2, p. 1, 2017.

[10] R. Ferguson, M. Sharples, and R. Beale, "Moocs 2030: a future for massive open online learning," *MOOCs and Open Education around the World*, pp. 315–326, 2015.

[11] P. A. Warusavitarana, K. Lokuge Dona, H. C. Piyathilake, D. D. Epitawela, and M. U. Edirisinghe, "Mooc: a higher education game changer in developing countries," *Rhetoric and Reality: Critical perspectives on educational technology. Proceedings ascilite Dunedin*, pp. 359–366, 2014.

[12] Z. W. Abas, "The glocalization of moocs in southeast asia," *Bonk, Lee, Reeves, & Reynolds (Ed.), MOOCs and open education: Around the world*, pp. 232–242, 2015.

[13] B. Venkataraman and A. Kanwar, "Changing the tune: Moocs for human development," *MOOCs and open education around the world*, pp. 206–217, 2015.

[14] A. Mutawa, "It is time to mooc and spoc in the gulf region," *Education and Information Technologies*, vol. 22, no. 4, pp. 1651–1671, 2017.

[15] M. H. Sallam, "A review of moocs in the arab world," *Creative Education*, vol. 8, no. 04, p. 564, 2017.

[16] Y. Pang, T. Wang, and N. Wang, "Mooc data from providers," in *2014 Enterprise Systems Conference*. IEEE, 2014, pp. 87–90.

[17] D. Jansen and R. Schuwer, "Institutional mooc strategies in europe," *Status Report Based on a Mapping Survey Conducted in October-December 2014*, 2015.

[18] M. Plourde, "Mooc poster v (3)," *Flickr, Abril*, vol. 4, 2013.

[19] G. Siemens, "Massive open online courses: Innovation in education?. commonwealth of learning," *Perspectives on Open and Distance Learning: Open Educational Resources: Innovation, Research and Practice*, p. 5, 2013.

[20] N. Spyropoulou, C. Pierrakeas, and A. Kameas, "Creating mooc guidelines based on best practices," *Edulearn14 Proceedings*, pp. 6981–6990, 2014.

[21] D. G. Glance, M. Forsey, and M. Riley, "The pedagogical foundations of massive open online courses," *First Monday*, vol. 18, no. 5, 2013.

[22] W. Admiraal, B. Huisman, and O. Pilli, "Assessment in massive open online courses." *Electronic Journal of E-learning*, vol. 13, no. 4, pp. 207–216, 2015.

[23] Udemy, "Certificate of completion," accessed: 2017-5-22. [Online]. Available: https://support.udemy.com/hc/en-us/articles/229603868-Certificate-of-Completion

[24] Coursera, "Course certificates," accessed: 2017-5-22. [Online]. Available: https://learner.coursera.help/hc/en-us/articles/208280196-Course-Certificates

[25] edX, "Certificates," accessed: 2017-5-22. [Online]. Available: https://support.edx.org/hc/en-us/sections/201019148-Certificates

[26] Futurelearn, "Certificate of achievement," accessed: 2017-5-22. [Online]. Available: https://www.futurelearn.com/proof-of-learning/certificate-of-achievement

[27] J. Sonwalkar and C. Maheshkar, "Moocs: a massive platform for collaborative learning in globalized way," *Journal of Management Research and Analysis*, vol. 2, no. 2, pp. 142–149, 2015.

[28] G. Siemens, "Connectivism: A learning theory for the digital age," *URL http://www.itdl.org/Journal/Jan_05/article01.htm*, 2004.

[29] O. Rodriguez, "The concept of openness behind c and x-moocs (massive open online courses)," *Open Praxis*, vol. 5, no. 1, pp. 67–73, 2013.

[30] A. T. Bates, *Teaching in a digital age: Guidelines for designing teaching and learning*, 2018.

[31] L. Pappano, "The year of the mooc," *The New York Times*, vol. 2, no. 12, p. 2012, 2012.

[32] R. Schuwer and B. Janssen, "Trends in business models for open edu-cational resources and open education," *Trend report: open educational resources 2013*, p. 60, 2013.

[33] T. Liyanagunawardena, A. Adams, and S. Williams, "Moocs: A systematic study of the published literature 2008-2012," *The international review of research in open and distributed learning*, vol. 14, no. 3, pp. 202–227, 2013.

[34] C. O. Rodriguez, "Moocs and the ai-stanford like courses: Two successful and distinct course formats for massive open online courses." *European Journal of Open, Distance and E-Learning*, 2012.

[35] T. Anderson and J. Dron, "Three generations of distance education pedagogy," *The International Review of Research in Open and Distributed Learning*, vol. 12, no. 3, pp. 80–97, 2011.

[36] M. Buhl and L. B. Andreasen, "Moocs beyond the c-and x-divide–the relevance of a social constructivist approach," in *MOOCs in Scandinavia Conference*, 2015.

[37] R. MUDZAMBA and E. De la Rey, "Moocs: A pedagogy of transformation in higher education," *Rhodes University Document, Department of Information Systems*, 2013.

[38] L. Yuan and S. Powell, "Moocs and open education: Implications for higher education," 2013.

[39] P. Adamopoulos, "What makes a great mooc? an interdisciplinary analysis of student retention in online courses," 2013.

[40] K. M. Alraimi, H. Zo, and A. P. Ciganek, "Understanding the moocs continuance: The role of openness and reputation," *Computers & Education*, vol. 80, pp. 28–38, 2015.

[41] B. G. Gameel, "Learner satisfaction with massive open online courses," *American Journal of Distance Education*, vol. 31, no. 2, pp. 98–111, 2017.

[42] B. K. Pursel, L. Zhang, K. W. Jablokow, G. Choi, and D. Velegol, "Understanding mooc students: motivations and behaviours indicative of mooc completion," *Journal of Computer Assisted Learning*, vol. 32, no. 3, pp. 202–217, 2016.

[43] P. De Barba, G. E. Kennedy, and M. Ainley, "The role of students' motivation and participation in predicting performance in a mooc," *Journal of Computer Assisted Learning*, vol. 32, no. 3, pp. 218–231, 2016.

[44] B. Huang and K. F. Hew, "Factors influencing learning and factors influencing persistence: A mixed-method study of mooc learners' motivation," in *Proceedings of the 2017 International Conference on Information System and Data Mining*. ACM, 2017, pp. 103–110.

[45] A. A. Mahrous and A. A. Ahmed, "A cross-cultural investigation of students' perceptions of the effectiveness of pedagogical tools: The middle east, the united kingdom, and the united states," *Journal of Studies in International Education*, vol. 14, no. 3, pp. 289–306, 2010.

[46] Cutestat. Web analysis for rwaq. Accessed: 2019-7-28. [Online]. Available: https://rwaq.org.cutestat.com

[47] A. Agarwal, "Moocs for the arab world," 2013. [Online]. Available: https://blog.edx.org/moocs-arab-world

[48] W. Sawahel, "New arab platform for moocs launched," *University World News, 30 May*, 2014.

[49] Edraak. عام إدراك الخامس. Accessed: 2019-7-4. [Online]. Available: https://blog.edraak.org

[50] Cutestat. Web analysis for edraak. Accessed: 2019-7-28. [Online]. Available: https://edraak.org.cutestat.com

[51] M. A. Jabbar, "Nadrus makes moocs arabic-friendly," 2016. [Online]. Available: https://www.wamda.com/2016/01/nadrus-makes-moocs-arabic-friendly

[52] edXPress, "A massive open online course (mooc) portal created exclusively for arab audiences will deliver vocational and employability skills to women, youth, persons with disabilities and citizens in rural communities," 2014. [Online]. Available: https://www.edx.org/press/saudi-arabia-edx-join-forces-bridge-gap

[53] Doroob. أسئلة عامة حول برنامج دروب. Accessed: 2019-3-12. [Online]. Available: https://www.doroob.sa/ar/individuals/faqs

[54] S. Ayad, "Saudi's rwaq team hope to create community of educators," 2015. [Online]. Available: https://www.wamda.com/2015/03/saudi-rwaq-introduces-new-education-platform

[55] Maharah. ماهو مهارة؟. Accessed: 2019-4-10. [Online]. Available: https://www.maharah.net/pages/about

[56] Zadi. حول زادي. Accessed: 2019-4-8. [Online]. Available: https://zadi.net/about

[57] M. Rahal, "The saudi women taking on the kingdom's education sector," 2017. [Online]. Available: https://www.wamda.com/2017/02/saudi-women-education-emkan

[58] Aanaab. ماهي أكاديمية أعناب؟. Accessed: 2019-4-8. [Online]. Available: https://www.aanaab.com/pages/about

[59] P. Rao, M. Komaraiah, and P. N. Reddy, "A case for moocs in indian higher education system," *Journal of Engineering Education Transformations*, vol. 29, no. 1, pp. 15–25, 2015.

[60] Y. Chengjie, "Challenges and changes of mooc to traditional classroom teaching mode," *Canadian Social Science*, vol. 11, no. 1, p. 135, 2015.

[61] B. Plangsorn, J. Na-Songkhla, and L. M. Luetkehans, "Undergraduate students' opinions with regard to ubiquitous mooc for enhancing cross-cultural competence." *World Journal on Educational Technology: Current Issues*, vol. 8, no. 3, p. 210, 2016.

[62] R. S. Adham and K. O. Lundqvist, "Moocs as a method of distance education in the arab world–a review paper," *European Journal of Open, Distance and E-Learning*, vol. 18, no. 1, pp. 123–138, 2015.

[63] M. Jemni and M. K. Khribi, "Toward empowering open and online education in the arab world through oer and moocs," in *Open education: from OERs to MOOCs*. Springer, 2017, pp. 73–100.

[64] R. S. Adham, "The use of avatars in gender-segregated blended learning and moocs in saudi arabia," Ph.D. dissertation, University of Reading., Reading, UK, 2017.

[65] R. Adham, P. Parslow, Y. Dimitriadi, and K. Ø. Lundqvist, "The use of avatars in gender segregated online learning within moocs in saudi arabia-a rwaq case study," *International Review of Research in Open and Distributed Learning*, vol. 19, no. 1, 2018.

[66] N. A. M. Hakami, "An investigation of the motivational factors influencing learners? intentions to continue using arabic moocs," Ph.D. dissertation, University of Southampton., Southampton, UK, 2018.

[67] M. Almuhanna, "Participants' perceptions of moocs in saudi arabia," Ph.D. dissertation, University of Sheffield., Sheffield, UK, 2018.

[68] P. G. Chadraba and R. D. O'Keefe, "Developing graduate marketing programs for economies in transition," *Journal of Marketing Education*, vol. 29, no. 3, pp. 218–222, 2007.

[69] A. Tubaishat, A. Bhatti, and E. El-Qawasmeh, "Ict experiences in two different middle eastern universities." *Issues in informing science & information technology*, vol. 3, 2006.

[70] J. A. Burt, "Impact of active learning on performance and motivation in female emirati students," *Learning & Teaching in Higher Education: Gulf Perspectives*, vol. 1, no. 4, pp. 1–15, 2004.

[71] N. Sonleitner and M. Khelifa, "Western-educated faculty challenges in a gulf classroom," *Learning and Teaching in Higher Education: Gulf Perspectives*, vol. 2, no. 1, pp. 1–21, 2005.

[72] A. Richards, "Higher education in egypt. policy research working papers series. education and employment." 1992.

[73] Journal of Educational Data Mining, "About the journal," accessed: 2018-4-8. [Online]. Available: https://jedm.educationaldatamining.org/index.php/JEDM

[74] G. Siemens and R. S. Baker, "Learning analytics and educational data mining: towards communication and collaboration," in *Proceedings of the 2nd international conference on learning analytics and knowledge*. ACM, 2012, pp. 252–254.

[75] M. Mayer, “Innovation at google: the physics of data,” in *PARC Forum*, 2009.

[76] K. R. Koedinger, R. S. Baker, K. Cunningham, A. Skogsholm, B. Leber, and J. Stamper, “A data repository for the edm community: The pslc datashop,” *Handbook of educational data mining*, vol. 43, pp. 43–56, 2010.

[77] R. S. Baker and P. S. Inventado, “Educational data mining and learning analytics,” in *Learning analytics*. Springer, 2014, pp. 61–75.

[78] C. Romero, S. Ventura, M. Pechenizkiy, and R. S. Baker, *Handbook of educational data mining*. CRC press, 2010.

[79] M. O. Pedro, R. Baker, A. Bowers, and N. Heffernan, “Predicting college enrollment from student interaction with an intelligent tutoring system in middle school,” in *Educational Data Mining 2013*, 2013.

[80] J. Cohen, “A coefficient of agreement for nominal scales,” *Educational and psychological measurement*, vol. 20, no. 1, pp. 37–46, 1960.

[81] J. Davis and M. Goadrich, “The relationship between precision-recall and roc curves,” in *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006, pp. 233–240.

[82] A. Merceron and K. Yacef, “Educational data mining: a case study.” in *AIED*, 2005, pp. 467–474.

[83] D. Ben-Naim, M. Bain, and N. Marcus, “A user-driven and data-driven approach for supporting teachers in reflection and adaptation of adaptive tutorials.” *International Working Group on Educational Data Mining*, 2009.

[84] I. Arroyo and B. P. Woolf, “Inferring learning and attitudes from a bayesian network of log file data.” in *AIED*, 2005, pp. 33–40.

[85] R. S. Baker, A. De Carvalho, J. Raspat, V. Aleven, A. T. Corbett, and K. R. Koedinger, “Educational software features that encourage and discourage “gaming the system”,” in *Proceedings of the 14th international conference on artificial intelligence in education*, 2009, pp. 475–482.

[86] D. Perera, J. Kay, I. Koprinska, K. Yacef, and O. R. Zaïane, “Clustering and sequential pattern mining of online collaborative learning data,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 6, pp. 759–772, 2008.

[87] D. H. Shanabrook, D. G. Cooper, B. P. Woolf, and I. Arroyo, “Identifying high-level student behavior using sequence-based motif discovery,” in *Educational Data Mining 2010*, 2010.

[88] P. Spirtes, C. N. Glymour, R. Scheines, D. Heckerman, C. Meek, G. Cooper, and T. Richardson, *Causation, prediction, and search*. MIT press, 2000.

[89] S. Fancsali, “Variable construction and causal discovery for cognitive tutor log data: Initial results,” in *Educational Data Mining 2012*, 2012.

[90] D. Rai and J. Beck, “Exploring user data from a game-like math tutor: a case study in causal modeling,” in *Educational Data Mining 2011*. Citeseer, 2010.

[91] L. Kaufman and P. J. Rousseeuw, “Finding groups in data: An introduction to cluster analysis–john wiley & sons,” *Inc., New York*, 1990.

[92] P. Kline, *An easy guide to factor analysis*. Routledge, 2014.

[93] N. Thai-Nghe, T. Horvath, and L. Schmidt-Thieme, “Context-aware factorization for personalized student’s task recommendation,” in *Proceedings of the international workshop on personalization approaches in learning environments*, vol. 732, 2011, pp. 13–18.

[94] H. Cen, K. Koedinger, and B. Junker, “Learning factors analysis–a general method for cognitive model evaluation and improvement,” in *International Conference on Intelligent Tutoring Systems*. Springer, 2006, pp. 164–175.

[95] R. S. Baker, S. M. Gowda, and A. T. Corbett, “Towards predicting future transfer of learning,” in *International Conference on Artificial Intelligence in Education*. Springer, 2011, pp. 23–30.

[96] J. MacQueen *et al.*, “Some methods for classification and analysis of multivariate observations,” in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 14, 1967, pp. 281–297.

[97] United Nations Development Programme, “Human development reports,” 2011. [Online]. Available: http://hdr.undp.org/en/media/HDR_2011_Statistical_Tables.xls.

[98] About human development. Accessed: 2018-10-30. [Online]. Available: http://hdr.undp.org/en/humandev/

[99] P. E. Black, "Manhattan distance," *Dictionary of Algorithms and Data Structures*, vol. 18, p. 2012, 2006.

[100] E. F. Krause, *Taxicab geometry: An adventure in non-Euclidean geometry*. Courier Corporation, 1986.

[101] P. R. Bahr, "The birdâ s eye view of community colleges: A behavioral typology of first-time students based on cluster analytic classification," *Research in Higher Education*, vol. 51, no. 8, pp. 724–749, 2010.

[102] V. B. Saenz, D. Hatch, B. E. Bukoski, S. Kim, K.-h. Lee, and P. Valdez, "Community college student engagement patterns: A typology revealed through exploratory cluster analysis," *Community College Review*, vol. 39, no. 3, pp. 235–267, 2011.

[103] D. Aloise, A. Deshpande, P. Hansen, and P. Popat, "Np-hardness of euclidean sum-of-squares clustering," *Machine learning*, vol. 75, no. 2, pp. 245–248, 2009.

[104] M. Mahajan, P. Nimbhorkar, and K. Varadarajan, "The planar k-means problem is np-hard," in *International Workshop on Algorithms and Computation*. Springer, 2009, pp. 274–285.

[105] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, pp. 832–837, 1956.

[106] W.-J. Wang, Y.-X. Tan, J.-H. Jiang, J.-Z. Lu, G.-L. Shen, and R.-Q. Yu, "Clustering based on kernel density estimation: nearest local maximum searching algorithm," *Chemometrics and intelligent laboratory systems*, vol. 72, no. 1, pp. 1–8, 2004.

[107] G. F. Jenks, "The data model concept in statistical mapping," *International yearbook of cartography*, vol. 7, pp. 186–190, 1967.

[108] H. Wang and M. Song, "Ckmeans. 1d. dp: optimal k-means clustering in one dimension by dynamic programming," *The R journal*, vol. 3, no. 2, p. 29, 2011.

[109] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53 – 65, 1987. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0377042787901257

[110] R. Bruning, G. Schraw, and M. Norby, *Cognitive Psychology and Instruction*. Pearson, 2011. [Online]. Available: https://books.google.co.uk/books?id=LPucSgAACAAJ

[111] L. Vygotsky, "Interaction between learning and development," *Readings on the development of children*, vol. 23, no. 3, pp. 34–41, 1978.

[112] V. I. Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals," in *Soviet physics doklady*, vol. 10, no. 8, 1966, pp. 707–710.

[113] M. Charrad, N. Ghazzali, V. Boiteau, and A. Niknafs, "Nbclust: An r package for determining the relevant number of clusters in a data set," *Journal of Statistical Software, Articles*, vol. 61, no. 6, pp. 1–36, 2014. [Online]. Available: https://www.jstatsoft.org/v061/i06

[114] P.-N. Tan, M. Steinbach, A. Karpatne, and V. Kumar, *Introduction to Data Mining (2Nd Edition)*, 2nd ed. Pearson, 2018.

[115] A. Field, *Discovering statistics using IBM SPSS statistics*. sage, 2013.

[116] J. Cohen, "Statistical power analysis for the behaviors science.(2nd)," *New Jersey: Laurence Erlbaum Associates, Publishers, Hillsdale*, 1988.

[117] Student, "The probable error of a mean," *Biometrika*, pp. 1–25, 1908.

[118] United Nations Development Programme, "Human development indices and indicators: 2018 statistical update," 2018. [Online]. Available: http://hdr.undp.org/sites/default/files/2018_human_development_statistical_update.pdf

[119] G. Stahl, T. D. Koschmann, and D. D. Suthers, *Computer-supported collaborative learning*. na, 2006.

[120] T. Staubitz, T. Pfeiffer, J. Renz, C. Willems, and C. Meinel, "Collaborative learning in a mooc environment," in *Proceedings of the 8th International Conference of Education, Research and Innovation*, 2015, pp. 8237–8246.

[121] G. Pison, A. Struyf, and P. J. Rousseeuw, "Displaying a clustering with clusplot," *Computational statistics & data analysis*, vol. 30, no. 4, pp. 381–392, 1999.

[122] I. Jolliffe, "Principal component analysis," *Technometrics*, vol. 45, no. 3, p. 276, 2003.

Appendices

A Human Development Index

Definition. *The Human Development Index (HDI) is an indicator that was devised in 1990 by economists Amartya Sen and Mahbub ul Haq. It is used by the United Nations Development Program to rank countries based on their health, knowledge and standard of living, into four levels of development [98].*

The index ranges from 0 and 1 and classifies countries into four levels of development: "Low" is from (0.000) to (0.554); "Medium" is from (0.555) to (0.699); "High" is from (0.700) to (0.799); and "Very High" is from (0.800) to (1.000). The HDI has been calculated using the following three dimensions [98]:

- **Health** is the measurement of life expectancy at birth.
- **Knowledge** is the average of the mean years of schooling and the expected years of schooling. The mean years of schooling is the average number of years of education received by people aged 25 years and older during their lifetime. The expected years of schooling is the average number of years of education received by people aged younger than 25 years.
- **Standard of living** is the measurement of gross national income (GNI) per capita.

Table 1 shows the three dimensions' indices and the maximum and minimum values of each dimension for 2011 [97]. The maximum values are the highest observations from

1980 until the recent year, and the minimum values are considered to be subsistence values.

Table 1: HDI dimensions indices with the maximum and minimum values from 2011

Dimension	Dimensions Indices	Indicators	Min-Max
Health	Life expectancy index (LEI)	Life expectancy at birth	20-83.4 years
Knowledge	Education index (EI)	Mean years of schooling Expected years of schooling	0-13.1 years 0-18 years
Standard of living	GNI index (GI)	GNI per capita	100-107,721 (PPP$)

The maximum and minimum values are used in the following three equations (1,4,5) to calculate the Dimensions Indices. Note that for GI, the natural logarithm (ln) is used to scale down the difference between the maximum and minimum values of the standard of living:

$$\mathbf{LEI} = \frac{\text{Actual value - Minimum value}}{\text{Maximum value - Minimum value}} \tag{1}$$

$$\text{Mean years of schooling } \mathbf{MYS} = \frac{\text{Actual value - Minimum value}}{\text{Maximum value - Minimum value}} \tag{2}$$

$$\text{Expected years of schooling } \mathbf{EYS} = \frac{\text{Actual value - Minimum value}}{\text{Maximum value - Minimum value}} \tag{3}$$

$$\mathbf{EI} = \frac{\text{MYS + EYS}}{2} \tag{4}$$

$$\mathbf{GI} = \frac{\ln(\text{Actual value}) - \ln(\text{Minimum value})}{\ln(\text{Maximum value}) - \ln(\text{Minimum value})} \tag{5}$$

The results of the Dimensions Indices equations are then used to calculate the HDI using equation 6. Based on the resulting index, countries are then classified as Low, Medium, High or Very High.

$$HDI = \sqrt[3]{LEI.EI.GI} \tag{6}$$

Example. *The actual values of a country's dimensions are collected by the United Nations Development Program. For this example, we assume the following values to be the actual values of country X:*

- *Life expectancy at birth is 80.2*
- *Mean years of schooling is 9.3 years*
- *Expected years of schooling 16.1 years*
- *Gross national income is 33,296 (PPP$)*

To calculate the HDI for country X, we use the above actual values and the maximum and minimum values from Table 1. Starting with the Life Expectancy Index (LEI), we calculate it as follows:

$$\boldsymbol{LEI} = \frac{\textit{Actual value - Minimum value}}{\textit{Maximum value - Minimum value}}$$

$$\boldsymbol{LEI} = \frac{80.2-20}{83.4-20}$$

$$\boldsymbol{LEI} = \frac{60.2}{63.4} = 0.949$$

We then calculate the Education Index (EI), which is the average of the mean years of schooling (MYS) and the expected years of schooling (EYS).

$$MYS = \frac{\textit{Actual value - Minimum value}}{\textit{Maximum value - Minimum value}}$$

$$MYS = \frac{9.3-0}{13.1-0} = 0.709$$

$EYS = \frac{\textit{Actual value - Minimum value}}{\textit{Maximum value - Minimum value}}$

$EYS = \frac{16.1-0}{18-0} = 0.894$

$\boldsymbol{EI} = \frac{MYS\ +\ EYS}{2}$

$\boldsymbol{EI} = \frac{0.709+0.894}{2} = 0.801$

Finally, we calculate the GNI Index (GI) as follows:

$\boldsymbol{GI} = \frac{\ln \textit{(Actual value)} - \ln \textit{(Minimum value)}}{\ln \textit{(Maximum value)} - \ln \textit{(Minimum value)}}$

$\boldsymbol{GI} = \frac{\ln(33{,}296) - \ln(100)}{\ln(107{,}721) - \ln(100)}$

$\boldsymbol{GI} = \frac{10.4-4.6}{11.5-4.6} = 0.840$

After calculating all Dimensions Indices, we can find the HDI for country X.

$HDI = \sqrt[3]{LEI.EI.GI}$

$HDI = \sqrt[3]{(0.949).(0.801).(0.840)}$

$HDI = 0.861$

Therefore, country X has a "Very High" level of development (between 1.000 and 0.800). In comparing this result with the United Nations Development Program's ranking in 2011, we find country X is in the 29th place.

B One-dimensional K-means Clustering Code in R

```
# set working directory
setwd("/Users/Desktop/One-dimensional␣K-means␣clustering")

# Read data file
Mydata <- read.csv("Edraak.csv")

# Combining the weekly scores into a data frame to form an engagement descriptions (ED)
ED <- data.frame(Mydata$One, Mydata$Two, Mydata$Three, Mydata$Four, Mydata$Five)

# Calculating similarities for the engagement descriptions (l_1 norm)
sim <- dist(ED, method = "manhattan")

# Validation test: (Clustering for K = 1:10)
ss <- rep(0,10)
for(i in 1:10){
       #Validation test: (Repeat clustering algorithm 100 times)
       clust_obj <- kmeans(sim, i, nstart = 100)
       #Calculating BSS/TSS ratio
       ss[i] <- clust_obj$betweenss/clust_obj$totss
}
summary(clust_obj)

# Plot the BSS/TSS ratio for K = 1:10 to choose the optimum K value
plot(c(1:10), ss, main="Choosing␣K-value", xlab="K-value", ylab="BSS/TSS␣ratio")

# Clustering with K = 3, and repeating the process 100 times
C <- kmeans(sim, 3, nstart = 100)

# Adding clustering result to the original data set
Mydata$cluster <- C$cluster

# Save to the data file
write.csv(Mydata, "Result.csv")
```

C Five-dimensional K-means Clustering Code in R

```
#set working directory
setwd("/Users/Shahad/Desktop/Five-dimensional K-means clustering")

# read data file for the "Equal-weighted" scoring system
Mydata1 <- read.csv("SS1.csv")

# read data file for the "Assessment-weighted" scoring system
Mydata2 <- read.csv("SS2.csv")

# read data file for the "Diligence-weighted" scoring system
Mydata3 <- read.csv("SS3.csv")

# read data file for the "Semi-diligence-weighted" scoring system
Mydata4 <- read.csv("SS4.csv")

# Combining the weekly scores, within each scoring system, into a data frame to form an engagement descriptions (ED)
ED1 <- data.frame(Mydata1$One, Mydata1$Two, Mydata1$Three, Mydata1$Four, Mydata1$Five)
ED2 <- data.frame(Mydata2$One, Mydata2$Two, Mydata2$Three, Mydata2$Four, Mydata2$Five)
ED3 <- data.frame(Mydata3$One, Mydata3$Two, Mydata3$Three, Mydata3$Four, Mydata3$Five)
ED4 <- data.frame(Mydata4$One, Mydata4$Two, Mydata4$Three, Mydata4$Four, Mydata4$Five)

# Validation test: (Clustering for K = 1:10 for the "Equal-weighted" scoring system)
ss1 <- rep(0,10)
for(i in 1:10){
       #Validation test: (Repeat clustering algorithm 100 times)
       clust_obj1 <- kmeans(ED1, i, nstart = 100)
       #Calculating BSS/TSS ratio
       ss1[i] <- clust_obj1$betweenss/clust_obj1$totss
}
summary(clust_obj1)

```

```
# Plot the BSS/TSS ratio for K = 1:10 to choose the optimum K value for the
    Equal-weighted scoring system
plot(c(1:10), ss1, main="Choosing K-value For The First Scoring System", xlab=
    "K-value", ylab="BSS/TSS ratio")

# Validation test: (Clustering for K = 1:10 for the "Assessment-weighted"
    scoring system)
ss2 <- rep(0,10)
for(i in 1:10){
       #Validation test: (Repeat clustering algorithm 100 times)
       clust_obj2 <- kmeans(ED2, i, nstart = 100)
       #Calculating BSS/TSS ratio
       ss2[i] <- clust_obj2$betweenss/clust_obj2$totss
}
summary(clust_obj2)

# Plot the BSS/TSS ratio for K = 1:10 to choose the optimum K value for the
    Assessment-weighted scoring system
plot(c(1:10), ss2, main="Choosing K-value For The Second Scoring System", xlab
    ="K-value", ylab="BSS/TSS ratio")

# Validation test: (Clustering for K = 1:10 for the "Diligence-weighted"
    scoring system)
ss3 <- rep(0,10)
for(i in 1:10){
       #Validation test: (Repeat clustering algorithm 100 times)
       clust_obj3 <- kmeans(ED3, i, nstart = 100)
       #Calculating BSS/TSS ratio
       ss3[i] <- clust_obj3$betweenss/clust_obj3$totss
}
summary(clust_obj3)

# Plot the BSS/TSS ratio for K = 1:10 to choose the optimum K value for the
    Diligence-weighted scoring system
plot(c(1:10), ss3, main="Choosing K-value For The Third Scoring System", xlab=
    "K-value", ylab="BSS/TSS ratio")

# Validation test: (Clustering for K = 1:10 for the "Semi-diligence-weighted"
    scoring system)
```

```
ss4 <- rep(0,10)
for(i in 1:10){
      #Validation test: (Repeat clustering algorithm 100 times)
      clust_obj4 <- kmeans(ED4, i, nstart = 100)
      #Calculating BSS/TSS ratio
      ss4[i] <- clust_obj4$betweenss/clust_obj4$totss
}
summary(clust_obj4)

# Plot the BSS/TSS ratio for K = 1:10 to choose the optimum K value for the Semi-diligence-weighted scoring system
plot(c(1:10), ss4, main="Choosing K-value For The Fourth Scoring System", xlab="K-value", ylab="BSS/TSS ratio")

# Validation test: (Using "NbClust" package to get the clusters indices, after running the K-means clustering algorithm)
install.packages("NbClust")
library(NbClust)

# Clustering Edraak data using the Equal-weighted scoring system
res<-NbClust(ED1, distance = "euclidean", min.nc=2, max.nc=10, method = "kmeans", index = "all")
# List all indices
ind <- res$All.index
# Plot the resulted clusters
plot(ind)

# Clustering Edraak data using the Assessment-weighted scoring system
res2<-NbClust(ED2, distance = "euclidean", min.nc=2, max.nc=10, method = "kmeans", index = "all")
# List all indices
ind2 <- res2$All.index
# Plot the resulted clusters
plot(ind2)

# Clustering Edraak data using the Diligence-weighted scoring system
res3<-NbClust(ED3, distance = "euclidean", min.nc=2, max.nc=10, method = "kmeans", index = "all")
# List all indices
```

```r
ind3 <- res3$All.index
# Plot the resulted clusters
plot(ind3)

# Clustering Edraak data using the Semi-diligence-weighted scoring system
res4<-NbClust(ED4, distance = "euclidean", min.nc=2, max.nc=10, method = "kmeans", index = "all")
# List all indices
ind4 <- res4$All.index
# Plot the resulted clusters
plot(ind4)

# Clustering with K=3, for the Equal-weighted scoring system, and repeating the process 100 times
C1 <- kmeans(ED1, 3, nstart = 100)
# Adding clustering result to original data set
Mydata1$clustering1 <- C1$cluster

# Clustering with K=3, for the Assessment-weighted scoring system, and repeating the process 100 times
C2 <- kmeans(ED2, 3, nstart = 100)
# Adding clustering result to original data set
Mydata2$clustering2 <- C2$cluster

# Clustering with K=3, for the Diligence-weighted scoring system, and repeating the process 100 times
C3 <- kmeans(ED3, 3, nstart = 100)
# Adding clustering result to original data set
Mydata3$clustering3 <- C3$cluster

# Clustering with K=3, for the Semi-diligence-weighted scoring system, and repeating the process 100 times
C4 <- kmeans(ED4, 3, nstart = 100)
# Adding clustering result to original data set
Mydata4$clustering4 <- C4$cluster

# Save to the data file
write.csv(Mydata1, "Result1.csv")
write.csv(Mydata2, "Result2.csv")
```

```
write.csv(Mydata3, "Result3.csv")
write.csv(Mydata4, "Result4.csv")
```

D Edraak learners' weekly interaction with MOOC: Using Kizilcec et al method

Table 2: Edraak learners' weekly interaction with video lectures
(Cluster 1: Sampling)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	61.03%	98.08%	99.65%	99.73%	99.47%
1	14.2%	1.12%	0.30%	0.26%	0.43%
2	9.77%	0.26%	0.04%	0%	0.08%
3	4.86%	0.13%	0%	0%	0%
4	2.95%	0.08%	0%	0%	0%
5	3.47%	0.04%	0%	0%	0%
6	3.69%	0.08%	0%	0%	0%
7		0%	0%		0%
8		0.13%	0%		0%
9		0.04%	0%		0%

* There are only 6 videos in week 2 and week 5

Table 3: Edraak learners' weekly interaction with assessments
(Cluster 1: Sampling)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	95.04%	100%	100%	100%	100%
Submitted	4.95%	0%	0%	0%	0%

Table 4: Edraak learners' weekly interaction with video lectures (Cluster 2: Disengaging)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	7.48%	8.26%	71.65%	93.3%	92.91%
1	0.39%	9.05%	8.26%	2.36%	3.14%
2	1.57%	11.41%	3.54%	0.78%	1.18%
3	4.72%	7.48%	2.36%	0.39%	0%
4	4.72%	9.84%	3.14%	0.78%	0.39%
5	25.59%	9.84%	1.18%	1.18%	0.39%
6	55.51%	7.48%	1.18%	1.18%	0%
7		4.72%	2.75%		0.78%
8		7.87%	2.75%		0.78%
9		24.01%	3.14%		0.39%

* There are only 6 videos in week 2 and week 5

Table 5: Edraak learners' weekly interaction with assessments (Cluster 2: Disengaging)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	6.69%	64.17%	93.3%	100%	99.21%
Submitted	93.3%	35.82%	6.69%	0%	0.78%

Table 6: Edraak learners' weekly interaction with video lectures (Cluster 3: Completing)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	36.66%	50%	55%	59.44%	63.33%
1	7.22%	4.44%	1.11%	2.77%	5%
2	3.88%	3.88%	1.11%	2.22%	3.33%
3	5.55%	0%	2.77%	3.88%	0.55%
4	0.55%	2.22%	1.11%	3.33%	0.55%
5	13.88%	1.11%	1.66%	7.77%	1.11%
6	32.22%	1.11%	1.11%	20.55%	1.66%
7		2.22%	3.33%		2.22%
8		7.77%	8.33%		5.55%
9		27.22%	24.44%		16.66%

* There are only 6 videos in week 2 and week 5

Table 7: Edraak learners' weekly interaction with assessments (Cluster 3: Completing)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	0.55%	0.55%	1.11%	7.22%	18.88%
Submitted	99.44%	99.44%	98.88%	92.77%	81.11%

E Edraak learners' weekly interaction with MOOC: Using Ferguson and Clow's method

Table 8: Edraak's learners (5-dimensional K-means clustering)

Scoring Systems	Cluster 1	Cluster 2	Cluster 3
Equal-weighted	2234	406	96
Assessment-weighted	2230	402	104
Diligence-weighted	2202	350	184
Semi-diligence-weighted	2203	348	185

E.1 Equal-weighted Scoring System

Table 9: Edraak learners' weekly interaction with video lectures (Cluster 1: Samplers)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	66.65%	98.52%	99.37%	99.55%	99.28%
1	15.26%	1.07%	0.49%	0.45%	0.54%
2	10.38%	0.27%	0.09%	0%	0.09%
3	5.01%	0.09%	0.04%	0%	0.04%
4	2.69%	0.04%	0%	0%	0%
5	0%	0%	0%	0%	0.04%
6	0%	0%	0%	0%	0%
7		0%	0%		0%
8		0%	0%		0%
9		0%	0%		0%

* There are only 6 videos in week 2 and week 5

Table 10: Edraak learners' weekly interaction with assessments (Cluster 1: Samplers)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	93.64%	95.7%	95.93%	96.2%	96.51%
Submitted	6.36%	4.3%	4.07%	3.8%	3.49%

Table 11: Edraak learners' weekly interaction with video lectures (Cluster 2: Mid-way Dropouts)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	0.25%	41.13%	87.19%	98.03%	95.07%
1	0%	7.88%	4.68%	0.74%	2.96%
2	0.99%	8.37%	2.46%	0.49%	1.48%
3	4.93%	4.93%	2.22%	0.49%	0%
4	4.68%	7.39%	2.22%	0%	0.25%
5	35.47%	6.65%	0.74%	0.25%	0%
6	53.69%	5.17%	0%	0%	0%
7		2.46%	0.49%		0.25%
8		4.68%	0%		0%
9		11.33%	0%		0%

* There are only 6 videos in week 2 and week 5

Table 12: Edraak learners' weekly interaction with assessments (Cluster 2: Mid-way Dropouts)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	27.09%	77.83%	92.61%	94.33%	95.32%
Submitted	72.91%	22.17%	7.39%	5.67%	4.68%

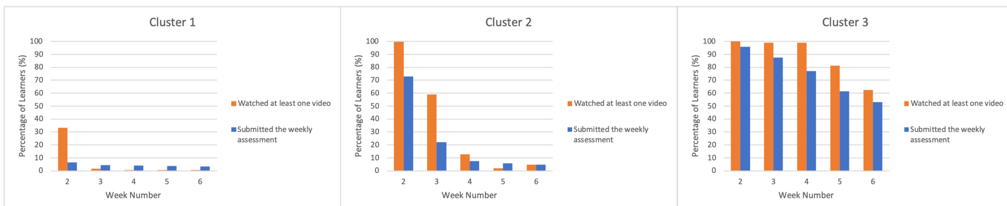
Table 13: Edraak learners' weekly interaction with video lectures (Cluster 3: Completers)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	0%	1.04%	1.04%	18.75%	37.5%
1	0%	1.04%	0%	4.17%	3.13%
2	0%	2.08%	0%	4.17%	3.13%
3	2.08%	0%	1.04%	6.25%	0%
4	2.08%	0%	1.04%	8.33%	1.04%
5	27.08%	1.04%	3.13%	16.67%	2.08%
6	68.75%	2.08%	5.21%	41.67%	3.13%
7		6.25%	11.46%		5.21%
8		18.75%	22.92%		12.5%
9		67.71%	54.17%		32.29%

* There are only 6 videos in week 2 and week 5

Table 14: Edraak learners' weekly interaction with assessments (Cluster 3: Completers)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	4.17%	12.5%	22.92%	38.54%	46.88%
Submitted	95.83%	87.5%	77.08%	61.46%	53.13%



(a) Samplers (b) Mid-way Dropouts (c) Completers

Figure -1: The weekly interactions with video lectures and assessments of the three clusters resulted from the Equal-weighted scoring system

E.2 Assessment-weighted Scoring System

Table 15: Edraak learners' weekly interaction with video lectures (Cluster 1: Samplers)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	66.77%	98.43%	99.42%	99.6%	99.33%
1	15.25%	1.12%	0.45%	0.4%	0.54%
2	10.13%	0.27%	0.09%	0%	0.09%
3	4.98%	0.13%	0.04%	0%	0%
4	2.87%	0.04%	0%	0%	0%
5	0%	0%	0%	0%	0.04%
6	0%	0%	0%	0%	0%
7		0%	0%		0%
8		0%	0%		0%
9		0%	0%		0%

* There are only 6 videos in week 2 and week 5

Table 16: Edraak learners' weekly interaction with assessments (Cluster 1: Samplers)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	94.04%	96.0%	96.23%	96.5%	96.77%
Submitted	5.96%	3.99%	3.77%	3.5%	3.23%

Table 17: Edraak learners' weekly interaction with video lectures (Cluster 2: Mid-way Dropouts)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	0.25%	43.03%	88.56%	98.26%	96.02%
1	0.25%	7.46%	4.98%	0.75%	2.49%
2	1.99%	8.46%	1.99%	0.25%	1.24%
3	4.73%	4.73%	1.74%	0.5%	0.25%
4	3.73%	6.72%	1.74%	0%	0%
5	35.32%	6.47%	0.75%	0.25%	0%
6	53.73%	4.98%	0.25%	0%	0%
7		2.24%	0%		0%
8		4.48%	0%		0%
9		11.44%	0%		0%

* There are only 6 videos in week 2 and week 5

Table 18: Edraak learners' weekly interaction with assessments (Cluster 2: Mid-way Dropouts)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	26.12%	78.11%	93.03%	94.28%	95.02%
Submitted	73.88%	21.89%	6.97%	5.72%	4.98%

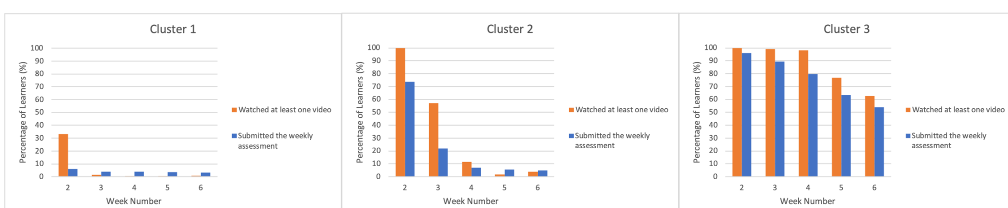
Table 19: Edraak learners' weekly interaction with video lectures (Cluster 3: Completers)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	0%	0.96%	1.92%	23.08%	37.5%
1	0%	1.92%	0%	4.81%	4.81%
2	1.92%	1.92%	1.92%	4.81%	3.85%
3	3.85%	0%	2.88%	5.77%	0%
4	1.92%	2.88%	2.88%	7.69%	1.92%
5	26.92%	1.92%	2.88%	15.38%	1.92%
6	65.38%	2.88%	3.85%	38.46%	2.88%
7		6.73%	12.5%		5.77%
8		18.27%	21.15%		11.54%
9		62.5%	50%		29.81%

* There are only 6 videos in week 2 and week 5

Table 20: Edraak learners' weekly interaction with assessments (Cluster 3: Completers)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	3.85%	10.58%	20.19%	36.54%	46.15%
Submitted	96.15%	89.42%	79.81%	63.46%	53.85%



(a) Samplers (b) Mid-way Dropouts (c) Completers

Figure -2: The weekly interactions with video lectures and assessments of the three clusters resulted from the Assessment-weighted scoring system

E.3 Diligence-weighted Scoring System

Table 21: Edraak learners' weekly interaction with video lectures (Cluster 1: Samplers)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	62.99%	97.59%	99.09%	99.55%	99.18%
1	14.44%	1.23%	0.59%	0.41%	0.59%
2	9.85%	0.36%	0.14%	0%	0.14%
3	5.09%	0.18%	0.05%	0%	0%
4	2.95%	0.14%	0.05%	0%	0%
5	2.91%	0.09%	0%	0%	0.05%
6	1.77%	0.09%	0%	0.05%	0%
7		0.09%	0%		0.05%
8		0.18%	0.09%		0%
9		0.05%	0%		0%

* There are only 6 videos in week 2 and week 5

Table 22: Edraak learners' weekly interaction with assessments (Cluster 1: Samplers)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	100%	100%	100%	100%	99.91%
Submitted	0%	0%	0%	0%	0.09%

Table 23: Edraak learners' weekly interaction with video lectures (Cluster 2: Mid-way Dropouts)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	10.57%	37.14%	84%	96%	96.29%
1	2.86%	6.29%	4.29%	0.86%	1.43%
2	3.14%	7.71%	2%	0.57%	1.14%
3	3.43%	5.14%	1.43%	0.86%	0%
4	4%	6.86%	2%	0.57%	0.29%
5	23.43%	6.86%	1.43%	0.86%	0%
6	52.57%	5.43%	0.57%	0.29%	0%
7		2.86%	1.43%		0.29%
8		5.14%	1.43%		0.57%
9		16.57%	1.43%		0%

* There are only 6 videos in week 2 and week 5

Table 24: Edraak learners' weekly interaction with assessments (Cluster 2: Mid-way Dropouts)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	0.57%	74.86%	96.57%	100 %	100%
Submitted	99.43%	25.14%	3.43%	0%	0%

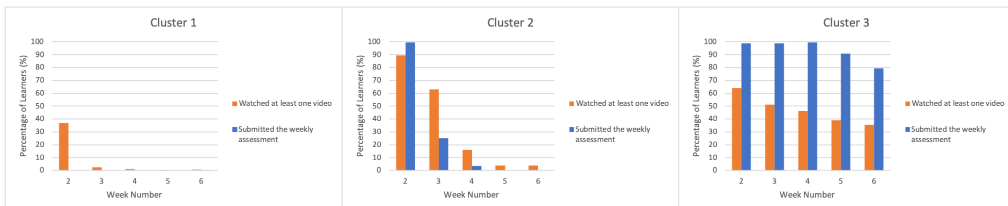
Table 25: Edraak learners' weekly interaction with video lectures (Cluster 3: Completers)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	35.87%	48.91%	53.8%	60.87%	64.67%
1	7.07%	4.35%	1.09%	2.72%	4.89%
2	4.35%	3.8%	1.09%	2.17%	2.17%
3	5.43%	0%	2.72%	2.72%	0.54%
4	1.09%	2.17%	1.09%	3.26%	0.54%
5	13.04%	1.09%	0.54%	7.61%	1.09%
6	33.15%	1.09%	1.63%	20.65%	1.63%
7		2.17%	4.35%		2.17%
8		8.15%	8.15%		5.43%
9		28.26%	25.54%		16.85%

* There are only 6 videos in week 2 and week 5

Table 26: Edraak learners' weekly interaction with assessments (Cluster 3: Completers)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	1.09%	1.09%	0.54%	9.24%	20.65%
Submitted	98.91%	98.91%	99.46%	90.76%	79.35%



(a) Samplers (b) Mid-way Dropouts (c) Completers

Figure -3: The weekly interactions with video lectures and assessments of the three clusters resulted from the Diligence-weighted scoring system

E.4 Semi-diligence-weighted Scoring System

Table 27: Edraak learners' weekly interaction with video lectures (Cluster 1: Samplers)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	63%	97.59%	99.09%	99.55%	99.18%
1	14.43%	1.23%	0.59%	0.41%	0.59%
2	9.85%	0.36%	0.14%	0%	0.14%
3	5.08%	0.18%	0.05%	0%	0%
4	2.95%	0.14%	0.05%	0%	0%
5	2.91%	0.09%	0%	0%	0.05%
6	1.77%	0.09%	0%	0.05%	0%
7		0.09%	0%		0.05%
8		0.18%	0.09%		0%
9		0.05%	0%		0%

* There are only 6 videos in week 2 and week 5

Table 28: Edraak learners' weekly interaction with assessments (Cluster 1: Samplers)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	100 %	99.95%	99.95%	99.95%	99.91%
Submitted	0%	0.05%	0.05%	0.05%	0.09%

Table 29: Edraak learners' weekly interaction with video lectures (Cluster 2: Mid-way Dropouts)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	10.63%	37.36%	84.48%	96.55%	96.84%
1	2.87%	6.32%	4.31%	0.86%	1.44%
2	3.16%	7.76%	2.01%	0.57%	0.86%
3	3.45%	5.17%	1.44%	0.57%	0%
4	4.02%	6.9%	2.01%	0.57%	0.29%
5	23.56%	6.9%	1.15%	0.86%	0%
6	52.3%	5.46%	0.57%	0%	0%
7		2.87%	1.44%		0.29%
8		5.17%	1.44%		0.29%
9		16.09%	1.15%		0%

* There are only 6 videos in week 2 and week 5

Table 30: Edraak learners' weekly interaction with assessments (Cluster 2: Mid-way Dropouts)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	0.57%	75%	96.55%	100%	100%
Submitted	99.43%	25%	3.45%	0%	0%

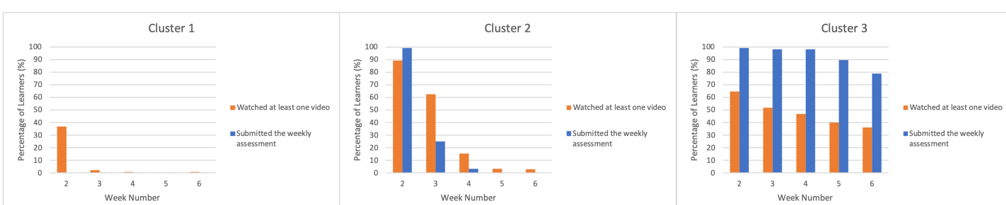
Table 31: Edraak learners' weekly interaction with video lectures (Cluster 3: Completers)

No. of Videos	Week 2*	Week 3	Week 4	Week 5*	Week 6
0	35.14%	48.11%	52.97%	60%	63.78%
1	7.03%	4.32%	1.08%	2.7%	4.86%
2	4.32%	3.78%	1.08%	2.16%	2.7%
3	5.41%	0%	2.7%	3.24%	0.54%
4	1.08%	2.16%	1.08%	3.24%	0.54%
5	12.97%	1.08%	1.08%	7.57%	1.08%
6	34.05%	1.08%	1.62%	21.08%	1.62%
7		2.16%	4.32%		2.16%
8		8.11%	8.11%		5.95%
9		29.19%	25.95%		16.76%

* There are only 6 videos in week 2 and week 5

Table 32: Edraak learners' weekly interaction with assessments (Cluster 3: Completers)

Assessments	Week 2	Week 3	Week 4	Week 5	Week 6
Not Submitted	0.54%	1.62%	1.62%	10.27%	21.08%
Submitted	99.46%	98.38%	98.38%	89.73%	78.92%



(a) Samplers (b) Mid-way Dropouts (c) Completers

Figure -4: The weekly interactions with video lectures and assessments of the three clusters resulted from the Semi-diligence-weighted scoring system

F Comparison between Edraak's engagement types: resulted from the Five-dimensional K-means clustering algorithm

To compare the demographics of the three engagement types of Edraak learners, the following statistical approach was used:

1. Is there a significant difference in the gender proportion among the engagement types? If so, what is the strength of this significant?
2. Is there a significant difference in the educational levels among the engagement types? If so, what is the strength of this significant?
3. Is there a significant difference in the distribution over countries with different HDI levels among the engagement types? If so, what is the strength of this significant?
4. Are there significant differences in the number of comments among the engagement types?
5. Are there significant differences in the exposure to the course materials among the engagement types?

F.1 Gender Proportions in the Engagement Types of Edraak

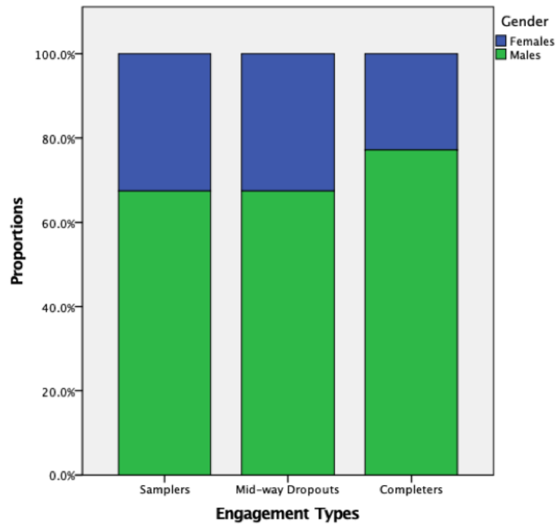


Figure -5: The gender proportions within the engagement types of Edraak learners

$X^2 = 7.487, df = 2, p = 0.024, V = .052$

Table 33: Crosstabulation of gender proportions and the engagement types of Edraak learners (Using Five-dimensional K-means clustering)

Gender	Engagement Types			Total
	Samplers	Mid-way Dropouts	Completers	
Males	1483	236	142	1861
% within gender	79.7%	12.7%	7.6%	100%
% within engagement type	67.3%	67.4%	77.2%	68%
Expected Count	1497.4	238.3	125.3	1861
Adjusted Residual	-1.5	-.3	2.7	
Females	716	114	42	872
% within gender	82.1%	13.1%	4.8%	100%
% within engagement type	32.5%	32.6%	22.8%	31.9%
Expected Count	701.6	111.7	58.7	872
Adjusted Residual	1.5	.3	-2.7	
Total	2199	350	184	2733
% within engagement type	100%	100%	100%	100%

F.2 Qualification Proportions in the Engagement Types of Edraak

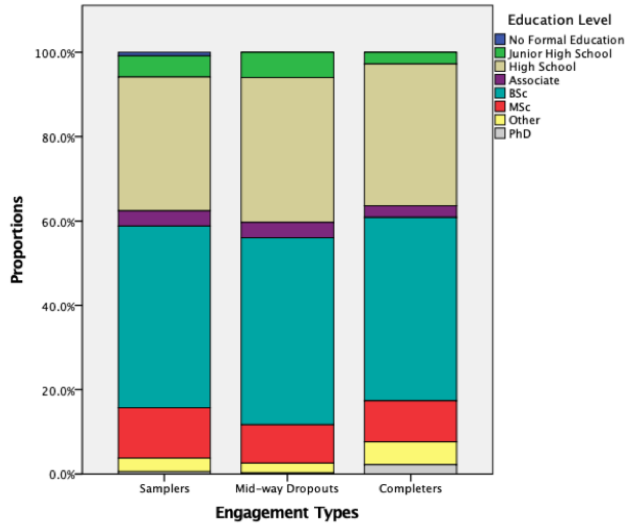


Figure -6: The qualification proportions within the engagement types of Edraak learners

$X^2 = 22.703, df = 14, p = .065$

Table 34: Crosstabulation of qualification proportions and the engagement types of Edraak learners (Using Five-dimensional K-means clustering)

Qualification	Engagement Types			Total
	Samplers	Mid-way Dropouts	Completers	
No formal education	19	0	0	19
% within qualification	100%	0%	0%	100%
% within engagement type	0.9%	0%	0%	0.7%
Expected Count	15.3	2.4	1.3	19
Adjusted Residual	2.2	-1.7	-1.2	
Junior high school	110	21	5	136
% within qualification	80.9%	15.4%	3.7%	100%
% within engagement type	5%	6%	2.7%	5%
Expected Count	109.5	17.4	9.1	136
Adjusted Residual	.1	.9	-1.5	
High school	699	120	62	881
% within qualification	79.3%	13.6%	7%	100%
% within engagement type	31.7%	34.3%	33.7%	32.2%

Continued on next page

Table 34 – continued from previous page

Qualification	Engagement Types			Total
	Samplers	Mid-way Dropouts	Completers	
Expected Count	709.1	112.7	59.2	881
Adjusted Residual	-1	.9	.4	
Associate	79	13	5	97
% within qualification	81.4%	13.4%	5.2%	100%
% within engagement type	3.6%	3.7%	2.7%	3.5%
Expected Count	78.1	12.4	6.5	97
Adjusted Residual	.2	.2	-.6	
BSc	949	155	80	1184
% within qualification	80.2%	13.1%	6.8%	100%
% within engagement type	43.1%	44.3%	43.5%	43.3%
Expected Count	952.9	151.5	79.6	1184
Adjusted Residual	-.4	.4	.1	
MSc	263	32	18	313
% within qualification	84%	10.2%	5.8%	100%
% within engagement type	11.9%	9.1%	9.8%	11.4%
Expected Count	251.9	40	21	313
Adjusted Residual	1.7	-1.4	-.7	
PhD	12	1	4	17
% within qualification	70.6%	5.9%	23.5%	100%
% within engagement type	0.5%	0.3%	2.2%	0.6%
Expected Count	13.7	2.2	1.1	17
Adjusted Residual	-1	-.9	2.8	
Other	71	8	10	89
% within qualification	79.8%	9%	11.2%	100%
% within engagement type	3.2%	2.3%	5.4%	3.3%
Expected Count	71.6	11.4	6	89
Adjusted Residual	-.2	-1.1	1.7	
Total	2202	350	184	2736
% within engagement type	100%	100%	100%	100%

F.3 Proportions of HDI levels in the Engagement Types of Edraak

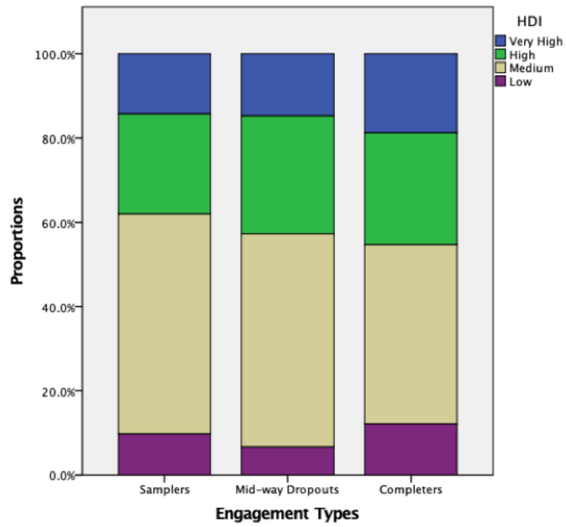


Figure -7: The proportions of HDI levels within the engagement types of Edraak learners

$X^2 = 12.346, df = 6, p = .055$

Table 35: Crosstabulation of the proportions of HDI level and the engagement types of Edraak learners (Using Five-dimensional K-means clustering)

HDI level	Engagement Types			Total
	Samplers	Mid-way Dropouts	Completers	
Very High	311	51	34	396
% within HDI level	78.5%	12.9%	8.6%	100%
% within engagement type	14.3%	14.7%	18.8%	14.7%
Expected Count	318.8	50.7	26.5	396
Adjusted Residual	-1.1	.0	1.6	
High	517	97	48	662
% within HDI level	78.1%	14.7%	7.3%	100%
% within engagement type	23.8%	28%	26.5%	24.5%
Expected Count	532.9	84.8	44.3	662
Adjusted Residual	-1.8	1.6	.7	
Medium	1135	175	77	1387
% within HDI level	81.8%	12.6%	5.6%	100%
% within engagement type	52.2%	50.6%	42.5%	51.3%
Expected Count	1116.5	177.6	92.9	1387
Adjusted Residual	1.8	-.3	-2.4	
Low	212	23	22	257
% within HDI level	82.5%	8.9%	8.6%	100%
% within engagement type	9.7%	6.6%	12.2%	9.5%
Expected Count	206.9	32.9	17.2	257
Adjusted Residual	.8	-1.9	1.3	
Total	2175	346	181	2702
% within engagement type	100%	100%	100%	100%

F.4 The Use of the Discussion Forum Among The Engagement Types of Edraak

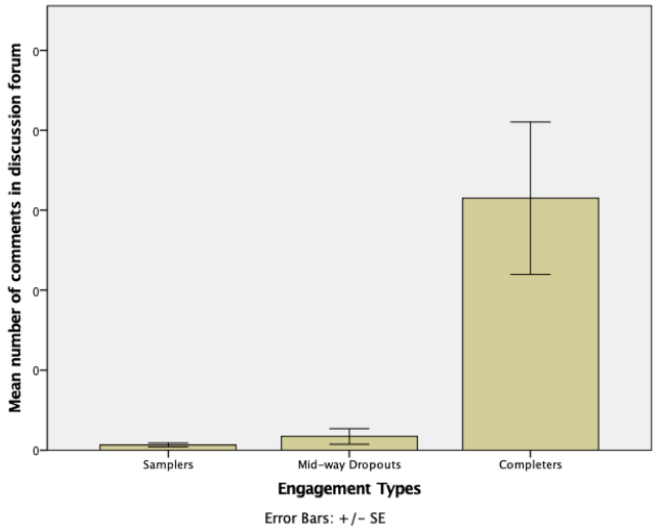


Figure -8: Comments distribution among the engagement types

The one-way ANOVA test showed a significant effect, $p < 0.001$, in the number of comments among the engagement types. The Games-Howell post-hoc test showed that the mean number of comments made by the "Completers" learners was significantly higher than the other two engagement types.

F.5 Exposure to MOOC Materials (Watching Video Lectures) Among the Engagement Types of Edraak

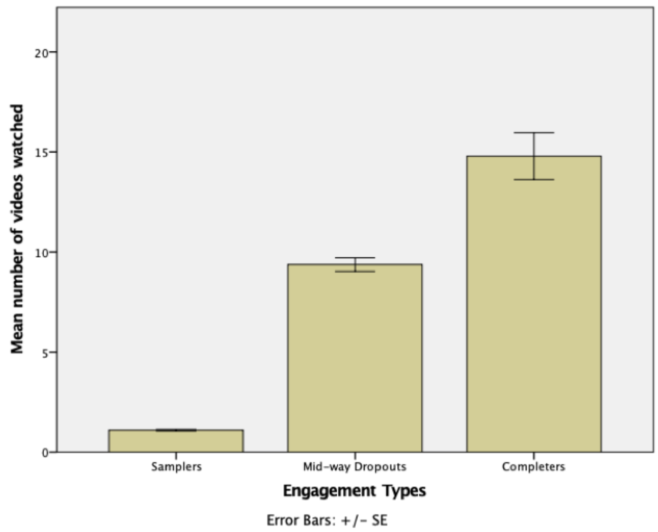


Figure -9: Engagement types interaction with video lectures

The one-way ANOVA test showed a significant effect among the compared groups, $p < 0.001$. The Games- Howell post-hoc test showed that the mean number of video lectures watched by the "Samplers" learners were significantly lower than the other engagement types, and "Mid-way Dropouts" learners were significantly lower than the "Completers".

G Comparison between Edraak's engagement types: One- versus Five-dimensional K-means clustering algorithm

G.1 Gender Proportions

Table 36: Crosstabulation of gender proportions and the Sampling/ Samplers engagement types of Edraak learners

		Sampling	Samplers	Total
Males	Count	1483	1546	3029
	Expected Count	1480.8	1548.2	3029
	% within Engagement Types	67.4%	67.2%	67.3%
	Adjusted Residual	.1	-.1	
Females	Count	716	753	1469
	Expected Count	718.2	750.8	1469
	% within Engagement Types	32.6%	32.8%	32.7%
	Adjusted Residual	-.1	.1	
Total	Count	2199	2299	4498
	% within Engagement Types	100%	100%	100%

$X^2 = .011, df = 1, p = .899$

Table 37: Crosstabulation of gender proportions and the Disengaging/ Mid-way Dropouts engagement types of Edraak learners

		Disengaging	Mid-way Dropouts	Total
Males	Count	177	236	413
	Expected Count	173.7	239.3	413
	% within Engagement Types	69.7%	67.4%	68.4%
	Adjusted Residual	.6	-.6	
Females	Count	77	114	191
	Expected Count	80.3	110.7	191
	% within Engagement Types	30.3%	32.6%	31.6%
	Adjusted Residual	-.6	.6	
Total	Count	254	350	604
	% within Engagement Types	100%	100%	100%

$X^2 = .250, df = 1, p = .595$

Table 38: Crosstabulation of gender proportions and the Completing/ Completers engagement types of Edraak learners

		Completing	Completers	Total
Males	Count	138	142	280
	Expected Count	138.5	141.5	280
	% within Engagement Types	76.7%	77.2%	76.9%
	Adjusted Residual	-.1	.1	
Females	Count	42	42	84
	Expected Count	41.5	42.5	84
	% within Engagement Types	23.3%	22.8%	23.1%
	Adjusted Residual	.1	-.1	
Total	Count	180	184	364
	% within Engagement Types	100%	100%	100%

$X^2 = .000, df = 1, p = 1$

G.2 Qualification Proportions

Table 39: Crosstabulation of qualification proportions and the Sampling/ Samplers engagement types of Edraak learners

		Sampling	Samplers	Total
No formal education	Count	19	19	38
	Expected Count	19.4	18.6	38
	% within Engagement Types	0.8%	0.9%	0.8%
	Adjusted Residual	-.1	.1	
Junior high school	Count	113	110	223
	Expected Count	114	109	223
	% within Engagement Types	4.9%	5%	5%
	Adjusted Residual	-.1	.1	
High school	Count	728	699	1427
	Expected Count	729.3	697.7	1427
	% within Engagement Types	31.6%	31.7%	31.7%
	Adjusted Residual	-.1	.1	
Associate	Count	81	79	160
	Expected Count	81.8	78.2	160
	% within Engagement Types	3.5%	3.6%	3.6%
	Adjusted Residual	-.1	.1	
BSc	Count	999	949	1948
	Expected Count	995.6	952.4	1948
	% within Engagement Types	43.4%	43.1%	43.3%
	Adjusted Residual	.2	-.2	
MSc	Count	277	263	540
	Expected Count	276	264	540
	% within Engagement Types	12%	11.9%	12%
	Adjusted Residual	.1	-.1	
PhD	Count	13	12	25
	Expected Count	12.8	12.2	25
	% within Engagement Types	0.6%	0.5%	0.6%
	Adjusted Residual	.1	-.1	
Other	Count	72	71	143
	Expected Count	72	71	143
	% within Engagement Types	3.1%	3.2%	3.2%
	Adjusted Residual	-.2	.2	
Total	Count	2302	2202	4504
	% within Engagement Types	100%	100%	100%

$X^2 = .128, df = 7, p = 1$

Table 40: Crosstabulation of qualification proportions and the Disengaging/ Mid-way Dropouts engagement types of Edraak learners

		Disengaging	Mid-way Dropouts	Total
Junior high school	Count	17	21	38
	Expected Count	16	22	38
	% within Engagement Types	6.7%	6%	6.3%
	Adjusted Residual	.3	-.3	
High school	Count	91	120	211
	Expected Count	88.7	122.3	211
	% within Engagement Types	35.8%	34.3%	34.9%
	Adjusted Residual	.4	-.4	
Associate	Count	11	13	24
	Expected Count	10.1	13.9	24
	% within Engagement Types	4.3%	3.7%	4%
	Adjusted Residual	.4	-.4	
BSc	Count	106	155	261
	Expected Count	109.8	151.2	261
	% within Engagement Types	41.7%	44.3%	43.2%
	Adjusted Residual	-.6	.6	
MSc	Count	20	32	52
	Expected Count	21.9	30.1	52
	% within Engagement Types	7.9%	9.1%	8.6%
	Adjusted Residual	-.5	.5	
PhD	Count	0	1	1
	Expected Count	.4	.6	1
	% within Engagement Types	0%	0.3%	0.2%
	Adjusted Residual	-.9	.9	
Other	Count	9	8	17
	Expected Count	7.1	9.9	17
	% within Engagement Types	3.5%	2.3%	2.8%
	Adjusted Residual	.9	-.9	
Total	Count	254	350	604
	% within Engagement Types	100%	100%	100%

$X^2 = 2.403, df = 6, p = .879$

Table 41: Crosstabulation of qualification proportions and the Completing/ Completers engagement types of Edraak learners

		Completing	Completers	Total
Junior high school	Count	6	5	11
	Expected Count	5.4	5.6	11
	% within Engagement Types	3.3%	2.7%	3%
	Adjusted Residual	.3	-.3	
High school	Count	62	62	124
	Expected Count	61.3	62.7	124
	% within Engagement Types	34.4%	33.7%	34.1%
	Adjusted Residual	.2	-.2	
Associate	Count	5	5	10
	Expected Count	4.9	5.1	10
	% within Engagement Types	2.8%	2.7%	2.7%
	Adjusted Residual	0	0	
BSc	Count	79	80	159
	Expected Count	78.6	80.4	159
	% within Engagement Types	43.9%	43.5%	43.7%
	Adjusted Residual	.1	-.1	
MSc	Count	16	18	34
	Expected Count	16.8	17.2	34
	% within Engagement Types	8.9%	9.8%	9.3%
	Adjusted Residual	-.3	.3	
PhD	Count	4	4	8
	Expected Count	4	4	8
	% within Engagement Types	2.2%	2.2%	2.2%
	Adjusted Residual	0	0	
Other	Count	8	10	18
	Expected Count	8.9	9.1	18
	% within Engagement Types	4.4%	5.4%	4.9%
	Adjusted Residual	-.4	.4	
Total	Count	180	184	364
	% within Engagement Types	100%	100%	100%

$X^2 = .393, df = 6, p = .999$

G.3 Proportions of HDI levels

Table 42: Crosstabulation of the proportions of HDI level and the Sampling/ Samplers engagement types of Edraak learners

		Sampling	Samplers	Total
Very High	Count	327	311	638
	Expected Count	326	312	638
	% within Engagement Types	14.4%	14.3%	14.3%
	Adjusted Residual	.1	-.1	
High	Count	543	517	1060
	Expected Count	541.7	518.3	1060
	% within Engagement Types	23.9%	23.8%	23.8%
	Adjusted Residual	.1	-.1	
Medium	Count	1185	1135	2320
	Expected Count	1185.6	1134.4	2320
	% within Engagement Types	52.1%	52.2%	52.2%
	Adjusted Residual	0	0	
Low	Count	218	212	430
	Expected Count	219.7	210.3	430
	% within Engagement Types	9.6%	9.7%	9.7%
	Adjusted Residual	-.2	.2	
Total	Count	2273	2175	4448
	% within Engagement Types	100%	100%	100%

$X^2 = .041, df = 3, p = .998$

Table 43: Crosstabulation of the proportions of HDI level and the Disengaging/ Mid-way Dropouts engagement types of Edraak learners

		Disengaging	Mid-way Dropouts	Total
Very High	Count	36	51	87
	Expected Count	36.6	50.4	87
	% within Engagement Types	14.3%	14.7%	14.6%
	Adjusted Residual	-.1	.1	
High	Count	74	97	171
	Expected Count	71.9	99.1	171
	% within Engagement Types	29.5%	28%	28.6%
	Adjusted Residual	.2	-.2	
Medium	Count	124	175	299
	Expected Count	125.7	173.3	299
	% within Engagement Types	49.4%	50.6%	50.1%
	Adjusted Residual	-.2	.1	
Low	Count	17	23	40
	Expected Count	16.8	23.2	40
	% within Engagement Types	6.8%	6.6%	6.7%
	Adjusted Residual	0	0	
Total	Count	251	346	597
	% within Engagement Types	100%	100%	100%

$X^2 = .166, df = 3, p = .983$

Table 44: Crosstabulation of the proportions of HDI level and the Completing/ Completers engagement types of Edraak learners

		Completing	Completers	Total
Very High	Count	33	34	67
	Expected Count	33.2	33.8	67
	% within Engagement Types	18.5%	18.8%	18.7%
	Adjusted Residual	-.1	.1	
High	Count	45	48	93
	Expected Count	46.1	46.9	93
	% within Engagement Types	25.3%	26.5%	25.9%
	Adjusted Residual	-.3	.3	
Medium	Count	78	77	155
	Expected Count	76.9	78.1	155
	% within Engagement Types	43.8%	42.5%	43.2%
	Adjusted Residual	.2	-.2	
Low	Count	22	22	44
	Expected Count	21.8	22.2	44
	% within Engagement Types	12.4%	12.2%	12.3%
	Adjusted Residual	.1	-.1	
Total	Count	178	181	359
	% within Engagement Types	100%	100%	100%

$X^2 = .093, df = 3, p = .993$